Interleukin-20

Ins A1

This application claims benefit under 35 U.S.C. § 119(e) of the filing date of copending U.S. Provisional Application Serial No. 60/052,870, filed on July 16, 1997, and copending U.S. Provisional Application Serial No. 60/060,140, filed on September 26, 1997. US Provisional Application Serial No. 60/060,140 is hereby incorporated herein by reference in its entirety.

Field of the Invention

The present invention relates to a novel human gene encoding a polypeptide which is a novel human cytokine. More specifically, isolated nucleic acid molecules are provided encoding a human polypeptide named Interleukin 20, hereinafter referred to as "IL-20". IL-20 polypeptides are also provided, as are vectors, host cells and recombinant methods for producing the same. Also provided are diagnostic methods for detecting disorders related to the immune system, and therapeutic methods for treating such disorders.

The invention further relates to screening methods for identifying agonists and antagonists of IL-20 activity.

Background of the Invention

Cytokines typically exert their respective biochemical and physiological effects by binding to specific receptor molecules. Receptor binding will then stimulate specific signal transduction pathways (Kishimoto, T., *et al., Cell* **76:**253-262 (1994). The specific interactions of cytokines with their receptors are often the primary regulators of a wide variety of cellular processes including activation, proliferation, and differentiation (Arai, K. -I, *et al., Ann. Rev. Biochem.* **59:**783-836 (1990); Paul, W. E. and Seder, R. A., *Cell* **76:**241-251 (1994)).

Human interleukin (IL)-17 was only recently identified. IL-17 is a 155 amino acid polypetide which was molecularly cloned from a CD4+ T-cell cDNA library (Yao, Z., *et al., J. Immunol.* **155:**5483-5486 (1995)). The IL-17 polypeptide contains an N-terminal signal peptide and contains approximately 72% identity at the amino acid level with a T-cell trophic herpesvirus saimiri (HVS) gene designated HVS13. High levels of IL-17 are secreted from CD4-positive primary peripheral blood leukocytes (PBL) upon stimulation (Yao, Z., *et al., Immunity* **3:**811-821 (1995)). Treatment of fibroblasts with IL-17, HVS13, or another murine homologue, designated CTLA8, activate signal transduction pathways and result in the stimulation of the NF-kB transcription factor family, the secretion of IL-6, and the costimulation of T-cell proliferation (Yao, Z., *et al., Immunity* **3:**811-821 (1995)).

An HVS13-Fc fusion protein was used to isolate a murine IL-17 receptor molecule which does not appear to belong to any of the previously described cytokine receptor families (Yao, Z., *et al., Immunity* **3:**811-821 (1995)). The murine IL-17 receptor (mIL-17R) is predicted to

encode a type I transmembrane protein of 864 amino acids with an apparent molecular mass of 97.8 kDa. mIL-17R is predicted to possess an N-terminal signal peptide with a cleavage site between alanine-31 and serine-32. The molecule also contains a 291 amino acid extracellular domain, a 21 amino acid transmembrane domain, and a 521 amino acid cytoplasmic tail. A soluble recombinant IL-17R molecule consisting of 323 amino acids of the extracellular domain of IL-17R fused to the Fc portion of human immunoglobulin IgG1 was able to significantly inhibit IL-17-induced IL-6 production by murine NIH-3T3 cells (*supra*).

Interestingly, the expression of the IL-17 gene is highly restricted. It is typically observed primarily in activated T-lymphocyte memory cells (Broxmeyer, H. *J. Exp. Med.* **183:**2411-2415 (1996); Fossiez, F., *et al., J. Exp. Med.* **183:**2593-2603 (1996)). Conversely, the IL-17 receptor appears to be expressed in a large number of cells and tissues (Rouvier, E., *et al., J. Immunol.* **150:**5445-5456 (1993); Yao, Z., *et al., J. Immunol.* **155:**5483-5486 (1995)). It remains to be seen, however, if IL-17 itself can play an autocrine role in the expression of IL-17. IL-17 has been implicated as a causitive agent in the expression of IL-6, IL-8, G-CSF, Prostaglandin E (PGE_2), and intracellular adhesion molecule (ICAM)-1 (Fossiez, F., *supra*; Yao, Z., *et al., Immunity* **3:**811-821 (1995)). Each of these molecules possesses highly relevent and potentially therapeutically valuable properties. For instance, IL-6 is involved in the regulation of hematopoietic stem and progenitor cell growth and expansion (Ikebuchi, K., *et al., Proc. Natl. Acad. Sci. USA* **84:**9035-9039 (1987); Gentile, P. and Broxmeyer, H. E. *Ann. N.Y. Acad. Sci. USA* **628:**74-83 (1991)). IL-8 exhibits a myelosuppressive activity for stem cells and immature subsets of myeloid progenitors (Broxmeyer, H. E., *et al., Ann. Hematol.* **71:**235-246 (1995); Daly, T. J., *et al., J. Biol. Chem.* **270:**23282-23292 (1995)). G-CSF acts both early and late to activate and stimulate hematopoiesis in general, and more specifically on neutrophil hematopoiesis, while PGE_2 enhances erythropoiesis, suppresses lymphopoiesis and myelopoiesis in general, and strongly suppresses monocytopoiesis (Broxmeyer, H. E. *Amer. J. Ped. Hematol./Oncol.* **14:**22-30 (1992); Broxmeyer, H. E. and Williams, D. E. *CRC Crit. Rev. Oncol./Hematol.* **8:**173-226 (1988)).

Thus, there is a need for polypeptides that function as immunoregulatory molecules and, thereby, function in the transfer of an extracellular signal ultimately to the nucleus of the cell, since disturbances of such regulation may be involved in disorders relating to cellular activation, hemostasis, angiogenesis, tumor metastasis, cellular migration and ovulation, as well as neurogenesis. Therefore, there is a need for identification and characterization of such human polypeptides which can play a role in detecting, preventing, ameliorating or correcting such disorders.

Summary of the Invention

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding at least a portion of the IL-20 polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 or the complete amino acid sequence encoded by the

cDNA clone (designated HTSGS30) deposited as plasmid DNA as ATCC Deposit Number 209232 on August 29, 1997. The nucleotide sequence determined by sequencing the deposited IL-20 clone, which is shown in Figure 1 (SEQ ID NO:1), contains an open reading frame encoding a complete polypeptide of 180 amino acid residues, including an initiation codon encoding an N-terminal methionine at nucleotide positions 45-47, and a predicted molecular weight of about 20.4 kDa. Nucleic acid molecules of the invention include those encoding the complete amino acid sequence excepting the N-terminal methionine shown in SEQ ID NO:2, or the complete amino acid sequence excepting the N-terminal methionine encoded by the human cDNA contained in HTSGS30, which molecules also can encode additional amino acids fused to the N-terminus of the IL-20 amino acid sequence.

The encoded polypeptide has a predicted leader sequence of 20 amino acids underlined in Figure 1; and the amino acid sequence of the predicted mature IL-20 protein is also shown in Figure 1, as amino acid residues 21-180 and as residues 1-160 in SEQ ID NO:2.

Thus, one aspect of the invention provides an isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding the IL-20 polypeptide having the complete amino acid sequence in SEQ ID NO:2 (i.e., positions -20 to 160 of SEQ ID NO:2); (b) a nucleotide sequence encoding the IL-20 polypeptide having the complete amino acid sequence in SEQ ID NO:2 excepting the N-terminal methionine (i.e., positions -19 to 160 of SEQ ID NO:2); (c) a nucleotide sequence encoding the predicted mature IL-20 polypeptide having the amino acid sequence at positions 1 to 160 in SEQ ID NO:2; (d) a nucleotide sequence encoding the IL-20 polypeptide having the complete amino acid sequence encoded by the human cDNA contained in HTSGS30; (e) a nucleotide sequence encoding the IL-20 polypeptide having the complete amino acid sequence excepting the N-terminal methionine encoded by the human cDNA contained in HTSGS30; (f) a nucleotide sequence encoding the mature IL-20 polypeptide having the amino acid sequence encoded by the human cDNA contained in HTSGS30; and, (g) a nucleotide sequence complementary to any of the nucleotide sequences in (a)-(f), above.

Further embodiments of the invention include isolated nucleic acid molecules that comprise a polynucleotide having a nucleotide sequence at least 90% identical to (that is to say, at most 10% different from), and more preferably at least 95%, 96%, 97%, 98% or 99% identical to (that is to say, at most 5%, 4%, 3%, 2% or 1% different from), any of the nucleotide sequences in (a), (b), (c), (d), (e), (f) or (g), above, or a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide in (a), (b), (c), (d), (e), (f) or (g), above. This polynucleotide which hybridizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues. An additional nucleic acid embodiment of the invention relates to an isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of a IL-20 polypeptide having an amino acid sequence in (a), (b), (c), (d), (e) or (f), above. A further nucleic acid embodiment of the invention relates to an isolated

nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of a IL-20 polypeptide having an amino acid sequence which contains at least one conservative amino acid substitution, but not more than 50 conservative amino acid substitutions, even more preferably, not more than 40 conservative amino acid substitutions, still more preferably not more than 30 conservative amino acid substitutions, and still even more preferably not more than 20 conservative amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a polynucleotide which encodes the amino acid sequence of a IL-20 polypeptide to have an amino acid sequence which contains not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 conservative amino acid substitutions.

The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of IL-20 polypeptides or peptides by recombinant techniques.

The invention further provides an isolated IL-20 polypeptide comprising an amino acid sequence selected from the group consisting of: (a) the amino acid sequence of the full-length IL-20 polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 (i.e., positions -20 to 160 of SEQ ID NO:2); (b) the amino acid sequence of the full-length IL-20 polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 excepting the N-terminal methionine (i.e., positions -19 to 160 of SEQ ID NO:2); (c) the amino acid sequence of the predicted mature IL-20 polypeptide having the amino acid sequence at positions 1 to 160 in SEQ ID NO:2; (d) the complete amino acid sequence encoded by the human cDNA contained in HTSGS30; (e) the complete amino acid sequence excepting the N-terminal methionine encoded by the human cDNA contained in HTSGS30; and (f) the complete amino acid sequence of the predicted mature IL-20 polypeptide encoded by the human cDNA contained in HTSGS30. The polypeptides of the present invention also include polypeptides having an amino acid sequence at least 80% identical to (that is to say, at most 20% different from), more preferably at least 90% identical to (that is to say, at most 10% different from), and still more preferably 95%, 96%, 97%, 98% or 99% identical to (that is to say, at most 5%, 4%, 3%, 2% or 1% different from) those described in (a), (b), (c), (d), (e) or (f), above, as well as polypeptides having an amino acid sequence with at least 90% similarity, and more preferably at least 95% similarity, to those above.

An additional embodiment of this aspect of the invention relates to a peptide or polypeptide which comprises the amino acid sequence of an epitope-bearing portion of a IL-20 polypeptide having an amino acid sequence described in (a), (b), (c), (d), (e) or (f), above. Peptides or polypeptides having the amino acid sequence of an epitope-bearing portion of a IL-20 polypeptide of the invention include portions of such polypeptides with at least six or seven, preferably at least nine, and more preferably at least about 30 amino acids to about 50 amino acids, although epitope-bearing polypeptides of any length up to and including the entire amino

acid sequence of a polypeptide of the invention described above also are included in the invention.

A further embodiment of the invention relates to a peptide or polypeptide which comprises the amino acid sequence of a IL-20 polypeptide having an amino acid sequence which contains at least one conservative amino acid substitution, but not more than 50 conservative amino acid substitutions, even more preferably, not more than 40 conservative amino acid substitutions, still more preferably not more than 30 conservative amino acid substitutions, and still even more preferably not more than 20 conservative amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a peptide or polypeptide to have an amino acid sequence which comprises the amino acid sequence of a IL-20 polypeptide, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 conservative amino acid substitutions.

In another embodiment, the invention provides an isolated antibody that binds specifically to a IL-20 polypeptide having an amino acid sequence described in (a), (b), (c), (d), (e) or (f) above. The invention further provides methods for isolating antibodies that bind specifically to a IL-20 polypeptide having an amino acid sequence as described herein. Such antibodies are useful diagnostically or therapeutically as described below.

The invention also provides for pharmaceutical compositions comprising IL-20 polypeptides, particularly human IL-20 polypeptides, which may be employed, for instance, to treat disorders relating to the proliferation or differentiation of T-cells, cellular activation, hemostasis, angiogenesis, tumor metastasis, cellular migration and ovulation, as well as neurogenesis. Methods of treating individuals in need of IL-20 polypeptides are also provided.

The invention further provides compositions comprising a IL-20 polynucleotide or an IL-20 polypeptide for administration to cells *in vitro*, to cells *ex vivo* and to cells *in vivo*, or to a multicellular organism. In certain particularly preferred embodiments of this aspect of the invention, the compositions comprise a IL-20 polynucleotide for expression of a IL-20 polypeptide in a host organism for treatment of disease. Particularly preferred in this regard is expression in a human patient for treatment of a dysfunction associated with aberrant endogenous activity of IL-20.

The present invention also provides a screening method for identifying compounds capable of enhancing or inhibiting a biological activity of the IL-20 polypeptide, which involves contacting a receptor which is enhanced by the IL-20 polypeptide with the candidate compound in the presence of a IL-20 polypeptide, assaying the IL-6 secretion or lymphocyte proliferation activity of the receptor in the presence of the candidate compound and of IL-20 polypeptide, and comparing the receptor activity to a standard level of activity, the standard being assayed when contact is made between the receptor and in the presence of the IL-20 polypeptide and the absence of the candidate compound In this assay, an increase in receptor activity over the standard indicates that the candidate compound is an agonist of IL-20 activity and a decrease in receptor activity compared to the standard indicates that the compound is an antagonist of IL-20 activity.

In another aspect, a screening assay for agonists and antagonists is provided which involves determining the effect a candidate compound has on IL-20 binding to a receptor. In particular, the method involves contacting the receptor with a IL-20 polypeptide and a candidate compound and determining whether IL-20 polypeptide binding to the receptor is increased or decreased due to the presence of the candidate compound. In this assay, an increase in binding of IL-20 over the standard binding indicates that the candidate compound is an agonist of IL-20 binding activity and a decrease in IL-20 binding compared to the standard indicates that the compound is an antagonist of IL-20 binding activity.

In yet another aspect, the IL-20 polypeptide may bind to a cell surface protein which also function as a viral receptor or coreceptor. Thus, IL-20, or agonists or antagonists thereof, may be used to regulate viral infectivity at the level of viral binding or interaction with the IL-20 receptor or coreceptor or during the process of viral internalization or entry into the cell.

It has been discovered that IL-20 is expressed not only in thymus, but also in thymus tumor and 12 week old whole human embryo. Therefore, nucleic acids of the invention are useful as hybridization probes for differential identification of the tissue(s) or cell type(s) present in a biological sample. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). In addition, for a number of disorders of the above tissues or cells, particularly of the immune, significantly higher or lower levels of IL-20 gene expression may be detected in certain tissues (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to a "standard" IL-20 gene expression level, i.e., the IL-20 expression level in healthy tissue from an individual not having the immune system disorder. Thus, the invention provides a diagnostic method useful during diagnosis of such a disorder, which involves: (a) assaying IL-20 gene expression level in cells or body fluid of an individual; (b) comparing the IL-20 gene expression level with a standard IL-20 gene expression level, whereby an increase or decrease in the assayed IL-20 gene expression level compared to the standard expression level is indicative of disorder in the immune system.

A further consequence of the observed thymus-restricted expression of endogenous IL-20 is that the IL-20 of the present invention may be useful in the regulation of the proliferation or differentiation of T-cells in general, for specific subsets of T-cells, for other immune cells in general, for other specific subsets of other immune cells, or any combination thereof. Thus, IL-20 of the present invention may be used therapeutically to treat disorders related to the immune system, including autoimmune and hematopoietic diseases or disorders, including AIDS, arthritis, or normal or abnormal cellular or systemic processes related to aging, and the like.

An additional aspect of the invention is related to a method for treating an individual in need of an increased level of IL-20 activity in the body comprising administering to such an individual a composition comprising a therapeutically effective amount of an isolated IL-20 polypeptide of the invention or an agonist thereof.

A still further aspect of the invention is related to a method for treating an individual in need of a decreased level of IL-20 activity in the body comprising, administering to such an individual a composition comprising a therapeutically effective amount of an IL-20 antagonist. Preferred antagonists for use in the present invention are IL-20-specific antibodies.

Brief Description of the Figures

Figure 1 shows the nucleotide sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) of IL-20. The predicted leader sequence of about 20 amino acids is underlined with a single underline. Note that the methionine residue at the beginning of the leader sequence in Figure 1 is shown in position number (positive) 1, whereas the leader positions in the corresponding sequence of SEQ ID NO:2 are designated with negative position numbers. Thus, the leader sequence positions 1 to 20 in Figure 1 correspond to positions -20 to -1 in SEQ ID NO:2.

A single potential asparagine-linked glycosylation site is marked in the amino acid sequence of IL-20 in Figure 1. The site is located from asparagine-75 through glutamic acid-78 in Figure 1 (N-75, S-76, S-77, E-78) [this sequence corresponds exactly to asparagine-55 through glutamic acid-58 in SEQ ID NO:2 (N-55, S-56, S-57, E-58)], and is marked with a bolded pound symbol (#) above the nucleotide sequence coupled with a bolded one letter abbreviation for the asparagine (**N**) in the amino acid sequence in Figure 1; that is, the actual asparagine residue which is potentially glycosylated is bolded in Figure 1.

Three potential Protein Kinase C (PKC) phosphorylation sites are also marked in Figure 1 with a bolded serine symbol (**S**) in the IL-20 amino acid sequence and an asterisk (*) above the first nucleotide encoding that serine residue in the IL-20 nucleotide sequence. The potential PKC phosphorylation sequences are found at the following locations in the IL-20 amino acid sequence: S-24 through K-26 (S-24, P-25, K-26); S-27 through R-29 (S-27, K-28, R-29); and S-93 through K-95 (S-93, N-94, K-95). A potential Casein Kinase II (CK2) phosphorylation site is also marked in Figure 1 with a bolded threonine symbol (**T**) in the IL-20 amino acid sequence and an asterisk (*) above the first nucleotide encoding the appropriate threonine residue in the IL-20 nucleotide sequence. The potential CK2 phosphorylation sequence is found at the following location in the IL-20 amino acid sequence: T-131 through E-134 (T-131, M-132, Q-133, K-134).

Regions of high identity between IL-20 and the closely related IL-21, IL-22, mouse IL-17, and viral IL-17 (an aligment of these sequences is presented in Figure 4) are delineated in Figure 1 with a double underline. These regions are not limiting and are labeled as Conserved Domain (CD)-I, CD-II, CD-III, and CD-IV in Figure 1.

Figure 2 shows the regions of identity between the amino acid sequences of the IL-20 protein and translation product of the human mRNA for IL-17 (SEQ ID NO:3), determined by the computer program Bestfit (Wisconsin Sequence Analysis Package, Version 8 for Unix,

Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711) using the default parameters.

Figure 3 shows an analysis of the IL-20 amino acid sequence. Alpha, beta, turn and coil regions; hydrophilicity and hydrophobicity; amphipathic regions; flexible regions; antigenic index and surface probability are shown. In the "Antigenic Index or Jameson-Wolf" graph, the positive peaks indicate locations of the highly antigenic regions of the IL-20 protein, i.e., regions from which epitope-bearing peptides of the invention can be obtained.

Detailed Description

The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding a IL-20 polypeptide having the amino acid sequence shown in SEQ ID NO:2, which was determined by sequencing a cloned cDNA. The nucleotide sequence shown in Figure 1 (SEQ ID NO:1) was obtained by sequencing the HTSGS30 clone, which was deposited on August 29, 1997 at the American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110, and given accession number ATCC 209232. The deposited clone is contained in the pBluescript SK(-) plasmid (Stratagene, La Jolla, CA). Clone HTSGS30 was also deposited with the ATCC as one of fifty plasmids pooled together in equimolar amounts. The pooled deposit was made on July 3, 1997 and was given ATCC Deposit No. 209138.

The IL-20 protein of the present invention shares sequence homology with the translation product of the human mRNA for IL-17 (Figure 2; SEQ ID NO:3). Human IL-17 is thought to be an important immunoregulatory molecule. The IL-17/IL-17 receptor complex activates NF-kB activity. NF-kB is a transcription factor known to regulate a large number of gene products involved in growth control. NF-kB-induced gene products include molecules involved in immune, inflammatory, or actute phase responses, such as immunoglobulin light chain, major histocompatibility complex (MHC), IL-2R a chain, and cytokines such as IL-1b, IL-6, and TNFa. NF-kB directly stimulates the HIV enhancer in T-cells and can itself be activated by different viral proteins with oncogenic potential, such as the hepatitis B virus HBX protein, EBV LMP1, and HTLV-1 Tax protein. The induction of NF-kB by Tax results in up-regulation of IL-2 and IL-2R and subsequently uncontrolled T-cell growth. IL-17 and HVS13, a gene product of HVS and a murine counterpart of IL-17, strongly induce IL-6 expression. IL-6 is a potent growth factor for myelomas, plasmacytomas, and hybridomas and is involved in the growth of Lennert's Lymphoma T-cells.

Nucleic Acid Molecules

Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer (such as the Model 373 from Applied Biosystems, Inc., Foster City, CA), and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence

determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

By "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U), where each thymidine deoxyribonucleotide (T) in the specified deoxyribonucleotide sequence is replaced by the ribonucleotide uridine (U).

Using the information provided herein, such as the nucleotide sequence in Figure 1 (SEQ ID NO:1), a nucleic acid molecule of the present invention encoding a IL-20 polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Illustrative of the invention, the nucleic acid molecule described in Figure 1 (SEQ ID NO:1) was discovered in a cDNA library derived from thymus. Additional clones of the same gene were also identified in thymus tumor and 12 week old whole human embryo cDNA libraries.

The determined nucleotide sequence of the IL-20 cDNA of Figure 1 (SEQ ID NO:1) contains an open reading frame encoding a protein of 180 amino acid residues, with an initiation codon at nucleotide positions 45-47 of the nucleotide sequence in Figure 1 (SEQ ID NO:1), and a deduced molecular weight of about 20.4 kDa. The amino acid sequence of the IL-20 protein shown in SEQ ID NO:2 is about 34.0 % identical to human mRNA for IL-17 (Figure 2; Yao, Z., *et al., J. Immunol.* **155:**5483-5486 (1995); GenBank Accession No. U32659).

Leader and Mature Sequences

The amino acid sequence of the complete IL-20 protein includes a leader sequence and a mature protein, as shown in SEQ ID NO:2. More in particular, the present invention provides nucleic acid molecules encoding a mature form of the IL-20 protein. Thus, according to the signal hypothesis, once export of the growing protein chain across the rough endoplasmic reticulum has been initiated, proteins secreted by mammalian cells have a signal or secretory leader sequence which is cleaved from the complete polypeptide to produce a secreted "mature" form of the protein. Most mammalian cells and even insect cells cleave secreted proteins with the same specificity. However, in some cases, cleavage of a secreted protein is not entirely uniform,

which results in two or more mature species of the protein. Further, it has long been known that the cleavage specificity of a secreted protein is ultimately determined by the primary structure of the complete protein, that is, it is inherent in the amino acid sequence of the polypeptide. Therefore, the present invention provides a nucleotide sequence encoding the mature IL-20 polypeptide having the amino acid sequence encoded by the human cDNA contained in clone HTSGS30. By the "mature IL-20 polypeptide having the amino acid sequence encoded by the human cDNA in HTSGS30" is meant the mature form(s) of the IL-20 protein produced by expression in a mammalian cell (e.g., COS cells, as described below) of the complete open reading frame encoded by the human DNA sequence of clone HTSGS30 which was deposited with the ATCC on two separate occasions.

In addition, methods for predicting whether a protein has a secretory leader as well as the cleavage point for that leader sequence are available. For instance, the method of McGeoch (*Virus Res.* **3**:271-286 (1985)) uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje (*Nucleic Acids Res.* **14**:4683-4690 (1986)) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2 where +1 indicates the amino terminus of the mature protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80% (von Heinje, *supra*). However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the complete IL-20 polypeptide was analyzed by a computer program "PSORT", available from Dr. Kenta Nakai of the Institute for Chemical Research, Kyoto University (Nakai, K. and Kanehisa, M. *Genomics* **14**:897-911 (1992)), which is an expert system for predicting the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. Thus, the computation analysis above predicted a single cleavage site within the complete amino acid sequence shown in SEQ ID NO:2.

As indicated, nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

In specific embodiments, the polynucleotides of this invention are less than 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb or 7.5 kb in length. In a further embodiment, polynucleotides of the invention comprise at least 15 contiguous nucleotides of IL-20 coding sequence, but do not comprise all or a portion of anyIL-20 intron. In another embodiment, the nucleic acid comprising IL-20 coding sequence does not contain coding sequences of a genomic flanking gene (i.e. 5' or 3' to the IL-20 coding sequence in the genome).

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment For example, recombinant DNA molecules

contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

Isolated nucleic acid molecules of the present invention include DNA molecules comprising an open reading frame (ORF) with an initiation codon at positions 45-47 of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1).

Also included are DNA molecules comprising the coding sequence for the predicted mature IL-20 protein shown at positions 1-160 of SEQ ID NO:2.

In addition, isolated nucleic acid molecules of the invention include DNA molecules which comprise a sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode the IL-20 protein. Of course, the genetic code and species-specific codon preferences are well known in the art. Thus, it would be routine for one skilled in the art to generate the degenerate variants described above, for instance, to optimize codon expression for a particular host (e.g., change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

In another aspect, the invention provides isolated nucleic acid molecules encoding the IL-20 polypeptide having an amino acid sequence encoded by the human cDNA contained in clone HTSGS30.

Preferably, this nucleic acid molecule will encode the mature polypeptide encoded by the above-described deposited cDNA clone.

The invention further provides an isolated nucleic acid molecule having the nucleotide sequence shown in Figure 1 (SEQ ID NO:1) or the nucleotide sequence of the IL-20 cDNA contained in the above-described deposited clone, or a nucleic acid molecule having a sequence complementary to one of the above sequences. Such isolated molecules, particularly DNA molecules, are useful as probes for gene mapping, by *in situ* hybridization with chromosomes, and for detecting expression of the IL-20 gene in human tissue, for instance, by Northern blot analysis.

The present invention is further directed to nucleic acid molecules encoding portions of the nucleotide sequences described herein as well as to fragments of the isolated nucleic acid molecules described herein. In particular, the invention provides a polynucleotide having a nucleotide sequence representing the portion of SEQ ID NO:1 which consists of positions 1-602 of SEQ ID NO:1.

In addition, the invention provides nucleic acid molecules having nucleotide sequences related to extensive portions of SEQ ID NO:1: HTYSK30Rb (SEQ ID NO:4). In one embodiment of the invention preferred polynucleotides do not consist of or comprise SEQ ID NO:4.

Further, the invention includes a polynucleotide comprising any portion of at least about 30 nucleotides, preferably at least about 50 nucleotides, of SEQ ID NO:1 from residue 60 to 599. Preferably, the invention includes a polynucleotide comprising any portion of at least about 30 nucleotides, preferably at least about 50 nucleotides, of SEQ ID NO:1 from residue 103 to 584. More preferably, the invention includes a polynucleotide comprising nucleotide residues 1-500, 25-525, 50-550, 75-575, 100-600, 125-625, 150-650, 175-675, 200-700, 103-595, 103-590, 103-585, 103-580, 103-575, 103-570, 103-565, 103-560, 103-555, 103-550, 103-545, 103-540, 103-535, 103-530, 103-525, 103-520, 103-515, and 103-510.

More generally, by a fragment of an isolated nucleic acid molecule having the nucleotide sequence of the deposited cDNA or the nucleotide sequence shown in Figure 1 (SEQ ID NO:1) is intended fragments at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length which are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments 50-300 nt in length are also useful according to the present invention as are fragments corresponding to most, if not all, of the nucleotide sequence of the deposited cDNA or as shown in Figure 1 (SEQ ID NO:1). By a fragment at least 20 nt in length, for example, is intended fragments which include 20 or more contiguous bases from the nucleotide sequence of the deposited cDNA or the nucleotide sequence as shown in Figure 1 (SEQ ID NO:1). Preferred nucleic acid fragments of the present invention include nucleic acid molecules encoding epitope-bearing portions of the IL-20 polypeptide as identified in Figure 3 and described in more detail below.

Also preferred in this aspect of the invention are fragments characterized by structural or functional attributes of IL-20. Preferred embodiments of the invention in this regard include fragments that comprise alpha-helix and alpha-helix forming regions ("alpha-regions"), beta-sheet and beta-sheet forming regions ("beta-regions"), turn and turn-forming regions ("turn-regions"), coil and coil-forming regions ("coil-regions"), hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions and high antigenic index regions of IL-20.

Certain preferred regions in these regards are set out in Figure 3, but may also be represented or identified by using a tabular representation of the data presented in Figure 3. The DNA*STAR computer algorithm used to generate Figure 3 (set on the original default parameters) will easily present the data in Figure 3 in such a tabular format. A tabular format of the data in Figure 3 may be used to easily determine specific boundaries of a preferred region.

The above-mentioned preferred regions set out in Figure 3 include, but are not limited to, regions of the aforementioned types identified by analysis of the amino acid sequence set out in Figure 1. As set out in Figure 3, such preferred regions include Garnier-Robson alpha-regions, beta-regions, turn-regions, and coil-regions, Chou-Fasman alpha-regions, beta-regions, and coil-regions, Kyte-Doolittle hydrophilic regions and hydrophobic regions, Eisenberg alpha- and beta-amphipathic regions, Karplus-Schulz flexible regions, Emini surface-forming regions and Jameson-Wolf regions of high antigenic index.

Among highly preferred fragments in this regard are those that comprise reigons of IL-20 that combine several structural features, such as several of the features set out above.

In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention described above, for instance, the cDNA clone HTSGS30 which is contained in ATCC Deposit No. 209232. By "stringent hybridization conditions" is intended overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (750 mM NaCl, 75 mM trisodium citrate) ~~(150 mM NaCl, 15 mM trisodium citrate)~~, 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

By a polynucleotide which hybridizes to a "portion" of a polynucleotide is intended a polynucleotide (either DNA or RNA) hybridizing to at least about 15 nucleotides (nt), and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably about 30-70 (e.g., 50) nt of the reference polynucleotide. These are useful as diagnostic probes and primers as discussed above and in more detail below.

By a portion of a polynucleotide of "at least 20 nt in length," for example, is intended 20 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide (e.g., the deposited cDNA or the nucleotide sequence as shown in Figure 1 (SEQ ID NO:1)). Of course, a polynucleotide which hybridizes only to a poly A sequence (such as the 3' terminal poly(A) tract of the IL-20 cDNA shown in Figure 1 (SEQ ID NO:1)), or to a complementary stretch of T (or U) residues, would not be included in a polynucleotide of the invention used to hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

As indicated, nucleic acid molecules of the present invention which encode an IL-20 polypeptide may include, but are not limited to those encoding the amino acid sequence of the mature polypeptide, by itself; and the coding sequence for the mature polypeptide and additional sequences, such as those encoding the about 20 amino acid leader or secretory sequence, such as a pre-, or pro- or prepro- protein sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences.

Also encoded by nucleic acids of the invention are the above protein sequences together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing, including splicing and polyadenylation signals, for example - ribosome binding and stability of mRNA; an additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities.

Thus, the sequence encoding the polypeptide may be fused to a marker sequence, such as a sequence encoding a peptide which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker amino acid sequence is a

hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described by Gentz and colleagues (*Proc. Natl. Acad. Sci. USA* **86:**821-824 (1989)), for instance, hexa-histidine provides for convenient purification of the fusion protein. The "HA" tag is another peptide useful for purification which corresponds to an epitope derived from the influenza hemagglutinin protein, which has been described by Wilson and coworkers (*Cell* **37:**767 (1984)). As discussed below, other such fusion proteins include the IL-20 fused to Fc at the N- or C-terminus.

Variant and Mutant Polynucleotides

The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the IL-20 protein. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism (*Genes II*, Lewin, B., ed., John Wiley & Sons, New York (1985)). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

Such variants include those produced by nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the IL-20 protein or portions thereof. Also especially preferred in this regard are conservative substitutions.

Polynucleotides encoding IL-20 muteins which contain amino acid mutations in regions of the IL-20 polypeptide sequence which exhibit a high degree of sequence identity with several closely related molecules (see Figure 4) have a high potential for possessing a change in an IL-20 biological activity. Such preferred embodiments may function as antagonists of innate IL-20 activities.

Most highly preferred are nucleic acid molecules encoding the mature protein having the amino acid sequence shown in SEQ ID NO:2 or the mature IL-20 amino acid sequence encoded by the deposited cDNA clone.

Thus, one aspect of the invention provides an isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding the IL-20 polypeptide having the complete amino acid sequence in SEQ ID NO:2 (i.e., positions -20 to 160 of SEQ ID NO:2); (b) a nucleotide sequence encoding the IL-20 polypeptide having the complete amino acid sequence in SEQ ID NO:2 excepting the N-terminal methionine (i.e., positions -19 to 160 of SEQ ID NO:2); (c) a nucleotide sequence encoding the predicted mature IL-20 polypeptide having the amino acid sequence at positions 1 to 160 in SEQ ID NO:2; (d) a nucleotide sequence encoding the IL-20 polypeptide having the

complete amino acid sequence encoded by the human cDNA in clone HTSGS30 (ATCC Deposit No. 209232); (e) a nucleotide sequence encoding the IL-20 polypeptide having the complete amino acid sequence excepting the N-terminal methionine encoded by the human cDNA in clone HTSGS30 (ATCC Deposit No. 209232); (f) a nucleotide sequence encoding the mature IL-20 polypeptide having the amino acid sequence encoded by the human cDNA clone HTSGS30 (ATCC Deposit No. 209232); and (g) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e) or (f), above.

Further embodiments of the invention include isolated nucleic acid molecules that comprise a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical, to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f) or (g), above, or a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide in (a), (b), (c), (d), (e), (f) or (g), above. This polynucleotide which hybridizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues. An additional nucleic acid embodiment of the invention relates to an isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of a IL-20 polypeptide having an amino acid sequence in (a), (b), (c), (d), (e) or (f), above. A further nucleic acid embodiment of the invention relates to an isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of a IL-20 polypeptide having an amino acid sequence which contains at least one conservative amino acid substitution, but not more than 50 conservative amino acid substitutions, even more preferably, not more than 40 conservative amino acid substitutions, still more preferably not more than 30 conservative amino acid substitutions, and still even more preferably not more than 20 conservative amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a polynucleotide which encodes the amino acid sequence of a IL-20 polypeptide to have an amino acid sequence which contains not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 conservative amino acid substitutions.

The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of IL-20 polypeptides or peptides by recombinant techniques.

In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly those of mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g. IL-20 coding sequence), and/or to include genetic material (e.g. heterologous polynucleotide sequences) that is operably associated with IL-20 polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous IL-20 polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g. promoter and/or enhancer) and endogenous IL-20

polynucleotide sequences via homologous recombination (see, e.g. U.S. Patent No. 5,641,670, issued June 24, 1997; Internation Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra, et al., Nature 342:435-438 (1989), the disclosures of each of which are hereby incorporated by reference in their entireties).

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a IL-20 polypeptide is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the IL-20 polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. The query sequence may be an entire sequence shown as SEQ ID NO:1 or any fragement specified as described herein.

As a practical matter, whether any particular nucleic acid molecule is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the nucleotide sequence shown in Figure 1 or to the nucleotides sequence of the deposited cDNA clone can be determined conventionally using known computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). Bestfit uses the local homology algorithm of Smith and Waterman to find the best segment of homology between two sequences (*Advances in Applied Mathematics* **2**:482-489 (1981)). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identiy are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group

Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the lenght of the subject nucleotide sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is becuase the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignement of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The present application is directed to nucleic acid molecules at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1) or to the nucleic acid sequence of the deposited cDNA, irrespective of whether they encode a polypeptide having IL-20 activity. This is because even where a particular nucleic acid molecule does not encode a polypeptide having IL-20 activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having IL-20 activity include, *inter alia*, (1) isolating the IL-20 gene or allelic variants thereof in a cDNA library; (2) *in situ* hybridization (e.g., "FISH") to metaphase chromosomal spreads to provide precise chromosomal location of the IL-20 gene, as described by Verma and colleagues (*Human Chromosomes: A Manual of Basic Techniques*, Pergamon

Press, New York (1988)); and Northern Blot analysis for detecting IL-20 mRNA expression in specific tissues.

Preferred, however, are nucleic acid molecules having sequences at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1) or to the nucleic acid sequence of the deposited cDNA which do, in fact, encode a polypeptide having IL-20 protein activity. By "a polypeptide having IL-20 activity" is intended polypeptides exhibiting activity similar, but not necessarily identical, to an activity of the mature IL-20 protein of the invention, as measured in a particular biological assay. For example, the IL-20 protein of the present invention modulates IL-6 secretion from NIH-3T3 cells. An *in vitro* ELISA assay which quantitates the amount of IL-6 secreted from cells in response to treatment with cytokines or the soluble extracellular domains of cytokine receptors has been described (Yao, Z., *et al.*, *Immunity* **3**:811-821 (1995)). Briefly, the assay involves plating the target cells at a density of approximately 5 x 10^6 cells/mL in a volume of 500 µL in the wells of a 24 well flat-bottomed culture plate (Costar). The cultures are then treated with various concentrations of the cytokine or the soluble extracellular domain of cytokine receptor in question The cells are then cultured for 24 hours at 37°C. At this time, 50 µL of supernatant is removed and assayed for the quantity of IL-6 essentially as described by the manufacturer (Genzyme, Boston, MA). IL-6 levels are then calculated by reference to a standard curve constructed with recombinant IL-17 cytokine. Such activity is useful for determining the level of IL-20-mediated IL-6 secretion.

IL-20 protein modulates immune system cell proliferation and differentiation in a dose-dependent manner in the above-described assay. Thus, "a polypeptide having IL-20 protein activity" includes polypeptides that also exhibit any of the same stimulatory activities in the above-described assays in a dose-dependent manner. Although the degree of dose-dependent activity need not be identical to that of the IL-20 protein, preferably, "a polypeptide having IL-20 protein activity" will exhibit substantially similar dose-dependence in a given activity as compared to the IL-20 protein (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity relative to the reference IL-20 protein).

Lymphocyte proliferation is another *in vitro* assay which may be performed to determine the activity of IL-20. For example, Yao and colleagues (*Immunity* **3**:811-821 (1995)) have recently described an *in vitro* assay for determining the effects of various cytokines and soluble cytokine receptors on the proliferation of murine leukocytes. Briefly, lymphoid organs are harvested aseptically, lymphocytes are isolated from the harvested organs, and the resulting collection of lymphoid cells are suspended in standard culture medium as described by Fanslow and coworkers (J. Immunol. 147:535-5540 (1991)). The lymphoid cell suspensions may then be divided into several different subclasses of lymphoid cells including splenic T-cells, lymph node B-cells, $CD4^+$ and $CD8^+$ T-cells, and mature adult thymocytes. For splenic T-cells, spleen cell suspensions (200 x 10^6 cells) are incubated with CD11b mAb and class II MHC mAb for 30 min at 4°C, loaded on a T-cell purification column (Pierce, Rockford, IL), and the T-cells eluted

according to the manufacturer's instructions. Using this method, purity of the resulting T-cell populations should be >95% $CD3^+$ and <1% $sIgM^+$. For purification of lymph node subsets, B-cells are removed from by adherence to tissue culture dishes previously coated with goat anti-mouse IgG (10μg/mL). Remaining cells were then incubated with anti-CD4 or anti-CD8 for 30 min at 4°C then washed and placed on tissue culture dishes previously coated with goat anti-rat IgG (20 μg/mL). After 45 min, nonadherent cells are removed and tested for purity by flow cytometry. CD4 and surface Ig-depleted cells should be >90% TCR-ab, $CD8^+$, whereas CD8 and surface Ig-depleted cells should be >95% TCR-ab, $CD4^+$. Finally, to enrich for mature adult thymocytes, cells are suspended at 10^8/mL in 10% anti-HSA and 10% low tox rabbit complement (Cedarlane, Ontario, Canada), incubated for 45 min at 37°C, and remaining viable cells isolated over Ficoll-Hypaque (Pharmacia, Piscataway, NJ). This procedure should yield between 90 and 95% $CD3^{hi}$ cells that are either $CD4^+8^-$ or $CD4^-8^+$.

To analyze the proliferative response of the above-described primary cell cultures, *in vitro* proliferation assays are set up in round bottom or flat bottom 96-well plates using 0.5-1.5 x 10^5 cells/well. For stimulation, T-cells are incubated with suboptimal concentrations (0.25-0.5 μg/mL) of Con A (Sigma, St. Louis, MO), PHA (0.25-0.5%; Difco, Detroit, MI), immobilized anti-CD3, or immobilized anti-TCR-ab. Anti-CD3 and anti-TCR-ab are immobilized for >2 hours at 37°C before the addition of effector cells. Incubations are done in the presence and absence of fixed CV-1/EBNA cells transfected with IL17RLP, muteins thereof, a control vector, or a control antigen such as rCD40L (Armitage, *et al., Nature* **357:**80 (1992)); Spriggs, *et al., J. Exp. Med.* **176:**1543 (1992)). Surface expression of CD40L is monitored by flow cytometry using a human CD40-Fc fusion protein. Cell cultures are pulsed overnight with [^{3}H]-thymidine (1 μCi/well) for the last 18 hours of a 3 day culture. Labeled cultures are then harvested on a 96-well Inotech harvester and radioactive counts detected using a scintillation counter.

Like other cytokines, IL-20 exhibits activity on leukocytes including for example monocytes, lymphocytes and neutrophils. For this reason IL-20 is active in directing the proliferation and differentiation of these cell types. Such activity is useful for immune enhancement or suppression, myeloprotection, stem cell mobilization, acute and chronic inflammatory control and treatment of leukemia. Assays for measuring such activity are well known in the art (Peters, *et al., Immun. Today* **17:**273 (1996); Young, *et al., J. Exp. Med.* **182:**1111 (1995); Caux, *et al., Nature* **390:**258 (1992); and Santiago-Schwarz, *et al., Adv. Exp. Med. Biol.* **378:**7 (1995).

Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 90%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequence of the deposited cDNA or the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1) will encode a polypeptide "having IL-20 protein activity." In fact, since degenerate variants of these nucleotide sequences all encode the same polypeptide, this will be clear to the skilled artisan even without performing the above described comparison assay. It will be further recognized in the art that, for such

nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having IL-20 protein activity. This is because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid), as further described below.

Vectors and Host Cells

The present invention also relates to vectors which include the isolated DNA molecules of the present invention, host cells which are genetically engineered with the recombinant vectors, and the production of IL-20 polypeptides or fragments thereof by recombinant techniques. The vector may be, for example, a phage, plasmid, viral or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged *in vitro* using an appropriate packaging cell line and then transduced into host cells.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac*, *trp*, *phoA* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293 and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9 (QIAGEN, Inc., *supra*); pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1, and pSG

(Stratagene); and pSVK3, pBPV, pMSG and pSVL (Pharmacia). Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals (for example, Davis, *et al., Basic Methods In Molecular Biology* (1986)).

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art. A preferred fusion protein comprises a heterologous region from immunoglobulin that is useful to stabilize and purify proteins. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is thoroughly advantageous for use in therapy and diagnosis and thus results, for example, in improved pharmacokinetic properties (EP-A 0232 262). On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified in the advantageous manner described. This is the case when Fc portion proves to be a hindrance to use in therapy and diagnosis, for example when the fusion protein is to be used as antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5 (Bennett, D., *et al., J. Molecular Recognition* **8**:52-58 (1995); Johanson, K., *et al., J. Biol. Chem.* **270**:9459-9471 (1995)).

The IL-20 protein can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Polypeptides of the present invention include: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production

procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Polypeptides and Fragments

The invention further provides an isolated IL-20 polypeptide having the amino acid sequence encoded by the deposited cDNA, or the amino acid sequence in SEQ ID NO:2, or a peptide or polypeptide comprising a portion of the above polypeptides.

Variant and Mutant Polypeptides

To improve or alter the characteristics of IL-20 polypeptides, protein engineering may be employed. Recombinant DNA technology known to those skilled in the art can be used to create novel mutant proteins or muteins including single or multiple amino acid substitutions, deletions, additions or fusion proteins. Such modified polypeptides can show, e.g., enhanced activity or increased stability. In addition, they may be purified in higher yields and show better solubility than the corresponding natural polypeptide, at least under certain purification and storage conditions.

N-Terminal and C-Terminal Deletion Mutants

For instance, for many proteins, including the extracellular domain of a membrane associated protein or the mature form(s) of a secreted protein, it is known in the art that one or more amino acids may be deleted from the N-terminus or C-terminus without substantial loss of biological function. For instance, Ron and colleagues (*J. Biol. Chem.*, **268:**2984-2988 (1993)) reported modified KGF proteins that had heparin binding activity even if 3, 8, or 27 N-terminal amino acid residues were missing. In the present case, since the protein of the invention is a member of the interleukin-17 polypeptide family, deletions of N-terminal amino acids up to the lysine at position 10 of SEQ ID NO:2 may retain some biological activity such as receptor binding or modulation of target cell activities. Polypeptides having further N-terminal deletions including the lys-10 residue in SEQ ID NO:2 would not be expected to retain such biological activities because this residue may be required for structural stability which is needed for receptor binding and signal transduction.

However, even if deletion of one or more amino acids from the N-terminus of a protein results in modification or loss of one or more biological functions of the protein, other biological activities may still be retained. Thus, the ability of the shortened protein to induce and/or bind to

antibodies which recognize the complete or mature form of the protein generally will be retained when less than the majority of the residues of the complete or mature form of the protein are removed from the N-terminus. Whether a particular polypeptide lacking N-terminal residues of a complete protein retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art.

Accordingly, the present invention further provides polypeptides having one or more residues deleted from the amino terminus of the amino acid sequence of the IL-20 shown in SEQ ID NO:2, up to the lysine residue at position number 10, and polynucleotides encoding such polypeptides. In particular, the present invention provides polypeptides comprising the amino acid sequence of residues n-160 of SEQ ID NO:2, where n is an integer in the range of -20 to 10, and 10 is the position of the first residue from the N-terminus of the complete IL-20 polypeptide (shown in SEQ ID NO:2) believed to be required for receptor binding activity of the IL-20 protein.

More in particular, the invention provides polynucleotides encoding polypeptides having the amino acid sequence of residues of -20-160, -19-160, -18-160, -17-160, -16-160, -15-160, -14-160, -13-160, -12-160, -11-160, -10-160, -9-160, -8-160, -7-160, -6-160, -5-160, -4-160, -3-160, -2-160, -1-160, 1-160, 2-160, 3-160, 4-160, 5-160, 6-160, 7-160, 8-160, 9-160, and 10-160, of SEQ ID NO:2. Polynucleotides encoding these polypeptides also are provided.

Similarly, many examples of biologically functional C-terminal deletion muteins are known. For instance, Interferon gamma shows up to ten times higher activities by deleting 8-10 amino acid residues from the carboxy terminus of the protein (Dobeli, *et al., J. Biotechnology* **7**:199-216 (1988)). In the present case, since the protein of the invention is a member of the interleukin-17 polypeptide family, deletions of C-terminal amino acids up to the cysteine at position 158 of SEQ ID NO:2 may retain some biological activity such as receptor binding or modulation of target cell activities, for chemokines. Polypeptides having further C-terminal deletions including the cysteine residue at position 158 of SEQ ID NO:2 would not be expected to retain such biological activities because this residue is likely to be required for forming a disulfide bridge to provide structural stability which is needed for receptor binding and signal transduction.

However, even if deletion of one or more amino acids from the C-terminus of a protein results in modification of loss of one or more biological functions of the protein, other biological activities may still be retained. Thus, the ability of the shortened protein to induce and/or bind to antibodies which recognize the complete or mature form of the protein generally will be retained when less than the majority of the residues of the complete or mature form of the protein are removed from the C-terminus. Whether a particular polypeptide lacking C-terminal residues of a complete protein retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art.

Accordingly, the present invention further provides polypeptides having one or more residues from the carboxy terminus of the amino acid sequence of the IL-20 shown in SEQ ID NO:2, up to the cysteine residue at position 158 of SEQ ID NO:2, and polynucleotides encoding

such polypeptides. In particular, the present invention provides polypeptides having the amino acid sequence of residues -20-m of the amino acid sequence in SEQ ID NO:2, where m is any integer in the range of 158 to 160, and residue 158 is the position of the first residue from the C-terminus of the complete IL-20 polypeptide (shown in SEQ ID NO:2) believed to be required for receptor binding activity of the IL-20 protein.

More in particular, the invention provides polynucleotides encoding polypeptides having the amino acid sequence of residues -20-158, -20-159, and -20-160 of SEQ ID NO:2. Polynucleotides encoding these polypeptides also are provided.

The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini, which may be described generally as having residues n-m of SEQ ID NO:2, where n and m are integers as described above.

Also included are a nucleotide sequence encoding a polypeptide consisting of a portion of the complete IL-20 amino acid sequence encoded by the human cDNA in clone HTSGS30 (ATCC Deposit No. 209232), where this portion excludes from 1 to about 30 amino acids from the amino terminus of the complete amino acid sequence encoded by the human cDNA in clone HTSGS30 (ATCC Deposit No. 209232), or from 1 to about 3 amino acids from the carboxy terminus, or any combination of the above amino terminal and carboxy terminal deletions, of the complete amino acid sequence encoded by the human cDNA in clone HTSGS30 (ATCC Deposit No. 209232). Polynucleotides encoding all of the above deletion mutant polypeptide forms also are provided.

As mentioned above, even if deletion of one or more amino acids from the N-terminus of a protein results in modification of loss of one or more biological functions of the protein, other biological activities may still be retained. Thus, the ability of the shortened IL-20 mutein to induce and/or bind to antibodies which recognize the complete or mature of the protein generally will be retained when less than the majority of the residues of the complete or mature protein are removed from the N-terminus. Whether a particular polypeptide lacking N-terminal residues of a complete protein retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art. It is not unlikely that an IL-20 mutein with a large number of deleted N-terminal amino acid residues may retain some biological or immungenic activities. In fact, peptides composed of as few as six IL-20 amino acid residues may often evoke an immune response.

Accordingly, the present invention further provides polypeptides having one or more residues deleted from the amino terminus of the IL-20 amino acid sequence shown in SEQ ID NO:2, up to the glycine residue at position number 175 and polynucleotides encoding such polypeptides. In particular, the present invention provides polypeptides comprising the amino acid sequence of residues n'-180 of Figure 1 (SEQ ID NO:2), where n' is an integer in the range of 2 to 175, and 176 is the position of the first residue from the N-terminus of the complete IL-20 polypeptide believed to be required for at least immunogenic activity of the IL-20 protein.

More in particular, the invention provides polynucleotides encoding polypeptides comprising, or alternatively consisting of, the amino acid sequence of residues of D-2 to F-180; W-3 to F-180; P-4 to F-180; H-5 to F-180; N-6 to F-180; L-7 to F-180; L-8 to F-180; F-9 to F-180; L-10 to F-180; L-11 to F-180; T-12 to F-180; I-13 to F-180; S-14 to F-180; I-15 to F-180; F-16 to F-180; L-17 to F-180; G-18 to F-180; L-19 to F-180; G-20 to F-180; Q-21 to F-180; P-22 to F-180; R-23 to F-180; S-24 to F-180; P-25 to F-180; K-26 to F-180; S-27 to F-180; K-28 to F-180; R-29 to F-180; K-30 to F-180; G-31 to F-180; Q-32 to F-180; G-33 to F-180; R-34 to F-180; P-35 to F-180; G-36 to F-180; P-37 to F-180; L-38 to F-180; A-39 to F-180; P-40 to F-180; G-41 to F-180; P-42 to F-180; H-43 to F-180; Q-44 to F-180; V-45 to F-180; P-46 to F-180; L-47 to F-180; D-48 to F-180; L-49 to F-180; V-50 to F-180; S-51 to F-180; R-52 to F-180; M-53 to F-180; K-54 to F-180; P-55 to F-180; Y-56 to F-180; A-57 to F-180; R-58 to F-180; M-59 to F-180; E-60 to F-180; E-61 to F-180; Y-62 to F-180; E-63 to F-180; R-64 to F-180; N-65 to F-180; I-66 to F-180; E-67 to F-180; E-68 to F-180; M-69 to F-180; V-70 to F-180; A-71 to F-180; Q-72 to F-180; L-73 to F-180; R-74 to F-180; N-75 to F-180; S-76 to F-180; S-77 to F-180; E-78 to F-180; L-79 to F-180; A-80 to F-180; Q-81 to F-180; R-82 to F-180; K-83 to F-180; C-84 to F-180; E-85 to F-180; V-86 to F-180; N-87 to F-180; L-88 to F-180; Q-89 to F-180; L-90 to F-180; W-91 to F-180; M-92 to F-180; S-93 to F-180; N-94 to F-180; K-95 to F-180; R-96 to F-180; S-97 to F-180; L-98 to F-180; S-99 to F-180; P-100 to F-180; W-101 to F-180; G-102 to F-180; Y-103 to F-180; S-104 to F-180; I-105 to F-180; N-106 to F-180; H-107 to F-180; D-108 to F-180; P-109 to F-180; S-110 to F-180; R-111 to F-180; I-112 to F-180; P-113 to F-180; V-114 to F-180; D-115 to F-180; L-116 to F-180; P-117 to F-180; E-118 to F-180; A-119 to F-180; R-120 to F-180; C-121 to F-180; L-122 to F-180; C-123 to F-180; L-124 to F-180; G-125 to F-180; C-126 to F-180; V-127 to F-180; N-128 to F-180; P-129 to F-180; F-130 to F-180; T-131 to F-180; M-132 to F-180; Q-133 to F-180; E-134 to F-180; D-135 to F-180; R-136 to F-180; S-137 to F-180; M-138 to F-180; V-139 to F-180; S-140 to F-180; V-141 to F-180; P-142 to F-180; V-143 to F-180; F-144 to F-180; S-145 to F-180; Q-146 to F-180; V-147 to F-180; P-148 to F-180; V-149 to F-180; R-150 to F-180; R-151 to F-180; R-152 to F-180; L-153 to F-180; C-154 to F-180; P-155 to F-180; P-156 to F-180; P-157 to F-180; P-158 to F-180; R-159 to F-180; T-160 to F-180; G-161 to F-180; P-162 to F-180; C-163 to F-180; R-164 to F-180; Q-165 to F-180; R-166 to F-180; A-167 to F-180; V-168 to F-180; M-169 to F-180; E-170 to F-180; T-171 to F-180; I-172 to F-180; A-173 to F-180; V-174 to F-180; and G-175 to F-180 of the IL-20 sequence shown in Figure 1 (which is identical to the sequence shown as SEQ ID NO:2, with the exception that the amino acid residues in Figure 1 are numbered consecutively from 1 through 180 from the N-terminus to the C-terminus, while the amino acid residues in SEQ ID NO:2 are numbered consecutively from -20 through 160 to reflect the position of the predicted signal peptide). Polynucleotides encoding these polypeptides are also encompassed by the invention.

Also as mentioned above, even if deletion of one or more amino acids from the C-terminus of a protein results in modification of loss of one or more biological functions of the

protein, other biological activities may still be retained. Thus, the ability of the shortened IL-20 mutein to induce and/or bind to antibodies which recognize the complete or mature of the protein generally will be retained when less than the majority of the residues of the complete or mature protein are removed from the C-terminus. Whether a particular polypeptide lacking C-terminal residues of a complete protein retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art. It is not unlikely that an IL-20 mutein with a large number of deleted C-terminal amino acid residues may retain some biological or immungenic activities. In fact, peptides composed of as few as six IL-20 amino acid residues may often evoke an immune response.

Accordingly, the present invention further provides polypeptides having one or more residues deleted from the carboxy terminus of the amino acid sequence of the IL-20 shown in SEQ ID NO:2, up to the asparagine residue at position number 6, and polynucleotides encoding such polypeptides. In particular, the present invention provides polypeptides comprising the amino acid sequence of residues 1-m' of SEQ ID NO:2, where m' is an integer in the range of 6 to 179, and 6 is the position of the first residue from the C-terminus of the complete IL-20 polypeptide believed to be required for at least immunogenic activity of the IL-20 protein.

More in particular, the invention provides polynucleotides encoding polypeptides comprising, or alternatively consisting of, the amino acid sequence of residues M-1 to I-179; M-1 to C-178; M-1 to T-177; M-1 to C-176; M-1 to G-175; M-1 to V-174; M-1 to A-173; M-1 to I-172; M-1 to T-171; M-1 to E-170; M-1 to M-169; M-1 to V-168; M-1 to A-167; M-1 to R-166; M-1 to Q-165; M-1 to R-164; M-1 to C-163; M-1 to P-162; M-1 to G-161; M-1 to T-160; M-1 to R-159; M-1 to P-158; M-1 to P-157; M-1 to P-156; M-1 to P-155; M-1 to C-154; M-1 to L-153; M-1 to R-152; M-1 to R-151; M-1 to R-150; M-1 to V-149; M-1 to P-148; M-1 to V-147; M-1 to Q-146; M-1 to S-145; M-1 to F-144; M-1 to V-143; M-1 to P-142; M-1 to V-141; M-1 to S-140; M-1 to V-139; M-1 to M-138; M-1 to S-137; M-1 to R-136; M-1 to D-135; M-1 to E-134; M-1 to Q-133; M-1 to M-132; M-1 to T-131; M-1 to F-130; M-1 to P-129; M-1 to N-128; M-1 to V-127; M-1 to C-126; M-1 to G-125; M-1 to L-124; M-1 to C-123; M-1 to L-122; M-1 to C-121; M-1 to R-120; M-1 to A-119; M-1 to E-118; M-1 to P-117; M-1 to L-116; M-1 to D-115; M-1 to V-114; M-1 to P-113; M-1 to I-112; M-1 to R-111; M-1 to S-110; M-1 to P-109; M-1 to D-108; M-1 to H-107; M-1 to N-106; M-1 to I-105; M-1 to S-104; M-1 to Y-103; M-1 to G-102; M-1 to W-101; M-1 to P-100; M-1 to S-99; M-1 to L-98; M-1 to S-97; M-1 to R-96; M-1 to K-95; M-1 to N-94; M-1 to S-93; M-1 to M-92; M-1 to W-91; M-1 to L-90; M-1 to Q-89; M-1 to L-88; M-1 to N-87; M-1 to V-86; M-1 to E-85; M-1 to C-84; M-1 to K-83; M-1 to R-82; M-1 to Q-81; M-1 to A-80; M-1 to L-79; M-1 to E-78; M-1 to S-77; M-1 to S-76; M-1 to N-75; M-1 to R-74; M-1 to L-73; M-1 to Q-72; M-1 to A-71; M-1 to V-70; M-1 to M-69; M-1 to E-68; M-1 to E-67; M-1 to I-66; M-1 to N-65; M-1 to R-64; M-1 to E-63; M-1 to Y-62; M-1 to E-61; M-1 to E-60; M-1 to M-59; M-1 to R-58; M-1 to A-57; M-1 to Y-56; M-1 to P-55; M-1 to K-54; M-1 to M-53; M-1 to R-52; M-1 to S-51; M-1 to V-50; M-1 to L-49; M-1 to D-48; M-1 to L-47; M-1 to P-46; M-1 to V-45; M-1 to Q-44; M-1 to H-43; M-1 to P-42; M-1 to G-41; M-1 to P-40; M-1 to A-39; M-1 to L-38;

M-1 to P-37; M-1 to G-36; M-1 to P-35; M-1 to R-34; M-1 to G-33; M-1 to Q-32; M-1 to G-31; M-1 to K-30; M-1 to R-29; M-1 to K-28; M-1 to S-27; M-1 to K-26; M-1 to P-25; M-1 to S-24; M-1 to R-23; M-1 to P-22; M-1 to Q-21; M-1 to G-20; M-1 to L-19; M-1 to G-18; M-1 to L-17; M-1 to F-16; M-1 to I-15; M-1 to S-14; M-1 to I-13; M-1 to T-12; M-1 to L-11; M-1 to L-10; M-1 to F-9; M-1 to L-8; M-1 to L-7; and M-1 to N-6 of the sequence of the IL-20 sequence shown in Figure 1 (which is identical to the sequence shown as SEQ ID NO:2, with the exception that the amino acid residues in Figure 1 are numbered consecutively from 1 through 180 from the N-terminus to the C-terminus, while the amino acid residues in SEQ ID NO:2 are numbered consecutively from -20 through 160 to reflect the position of the predicted signal peptide). Polynucleotides encoding these polypeptides also are provided.

The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini of an IL-20 polypeptide, which may be described generally as having residues n'-m' of Figure 1 (SEQ ID NO:2), where n' and m' are integers as described above.

A specific embodiment of the invention is an IL-20 deletion mutation expression construct for expression in E. coli. More specifically, this IL-20 deletion mutation expression construct for expression in E. coli will contain a nucleic acid insert which encodes amino acids glutamine-21 through phenlyalanine-180 of the IL-20 polypeptide sequence shown in Figure 1 (or amino acids glutamine-1 through phenlyalanine-160 of the IL-20 polypeptide sequence shown in SEQ ID NO:2).

Additional embodiments of the invention include the complete polypeptide sequence shown as SEQ ID NO:11 and specific portions thereof. This polypeptide sequence was obtained by translation of a polynucleotide sequence assembled by methodology other than that used to construct the sequence shown as SEQ ID NO:1. Preferred are polypeptides comprising residues 1-127, 2-127, and 62-127 all of SEQ ID NO:11. Also preferred are isolated polypeptides comprising at least ten contigous amino acids, more preferrably at least 30 contigous amino acids, from SEQ ID NO:11. Polynucleotides encoding the such polypeptide sequences are also provided.

Other Mutants

In addition to terminal deletion forms of the protein discussed above, it also will be recognized by one of ordinary skill in the art that some amino acid sequences of the IL-20 polypeptide can be varied without significant effect of the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine activity.

Thus, the invention further includes variations of the IL-20 polypeptide which show substantial IL-20 polypeptide activity or which include regions of IL-20 protein such as the protein portions discussed below. Such mutants include deletions, insertions, inversions, repeats, and type substitutions selected according to general rules known in the art so as have

little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided wherein the authors indicate that there are two main approaches for studying the tolerance of an amino acid sequence to change (Bowie, J. U., *et al., Science* **247:**1306-1310 (1990)),. The first method relies on the process of evolution, in which mutations are either accepted or rejected by natural selection. The second approach uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene and selections or screens to identify sequences that maintain functionality.

As the authors state, these studies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at a certain position of the protein. For example, most buried amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Other such phenotypically silent substitutions are described by Bowie and coworkers (*supra*) and the references cited therein. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

Thus, the fragment, derivative or analog of the polypeptide of SEQ ID NO:2, or that encoded by the deposited cDNA, may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the above form of the polypeptide, such as an IgG Fc fusion region peptide or leader or secretory sequence or a sequence which is employed for purification of the above form of the polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

Thus, the IL-20 of the present invention may include one or more amino acid substitutions, deletions or additions, either from natural mutations or human manipulation. As indicated, changes are preferably of a minor nature, such as conservative amino acid substitutions that do not significantly affect the folding or activity of the protein (see Table 1).

TABLE 1. Conservative Amino Acid Substitutions.

Aromatic	Phenylalanine
	Tryptophan
	Tyrosine
Hydrophobic	Leucine
	Isoleucine
	Valine
Polar	Glutamine
	Asparagine
Basic	Arginine
	Lysine
	Histidine
Acidic	Aspartic Acid
	Glutamic Acid
Small	Alanine
	Serine
	Threonine
	Methionine
	Glycine

In specific embodiments, the number of substitutions, deletions or additions in the amino acid sequence of Figure 1 and/or any of the polypeptide fragments described herein is 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12 ,11, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 or 30-20, 20-10, 20-15, 15-10, 10-5 or 1-5.

Amino acids in the IL-20 protein of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, *Science* **244:**1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as receptor binding or *in vitro* proliferative activity.

Of special interest are substitutions of charged amino acids with other charged or neutral amino acids which may produce proteins with highly desirable improved characteristics, such as less aggregation. Aggregation may not only reduce activity but also be problematic when preparing pharmaceutical formulations, because aggregates can be immunogenic (Pinckard, *et*

al., Clin. Exp. Immunol. **2:**331-340 (1967); Robbins, *et al., Diabetes* **36:**838-845 (1987); Cleland, *et al., Crit. Rev. Therapeutic Drug Carrier Systems* **10:**307-377 (1993)).

Replacement of amino acids can also change the selectivity of the binding of a ligand to cell surface receptors (for example, Ostade, *et al., Nature* **361:**266-268 (1993)) describes certain mutations resulting in selective binding of TNF-α to only one of the two known types of TNF receptors. Sites that are critical for ligand-receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith, *et al., J. Mol. Biol.* **224:**899-904 (1992); de Vos, *et al. Science* **255:**306-312 (1992)).

Since IL-20 is a member of the cytokine-related protein family, to modulate rather than completely eliminate biological activities of IL-20 preferably mutations are made in sequences encoding amino acids in the IL-20 conserved domain, i.e., in positions -14 to 158 of SEQ ID NO:2, more preferably in residues 76 to 158 of SEQ ID NO:2, and even more preferably, in residues within this region which are not conserved in all cytokines and cytokine-like molecules. Also forming part of the present invention are isolated polynucleotides comprising nucleic acid sequences which encode the above IL-20 mutants.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of the IL-20 polypeptide can be substantially purified by the one-step method described by Smith and Johnson (*Gene* **67:**31-40 (1988)). Polypeptides of the invention also can be purified from natural or recombinant sources using anti-IL-20 antibodies of the invention in methods which are well known in the art of protein purification.

The invention further provides an isolated IL-20 polypeptide comprising an amino acid sequence selected from the group consisting of: (a) the amino acid sequence of the full-length IL-20 polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 (i.e., positions -20 to 160 of SEQ ID NO:2); (b) the amino acid sequence of the full-length IL-20 polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 excepting the N-terminal methionine (i.e., positions -19 to 160 of SEQ ID NO:2); (c) the amino acid sequence of the predicted mature IL-20 polypeptide having the amino acid sequence at positions 1 to 160 in SEQ ID NO:2; (d) the complete amino acid sequence encoded by the human cDNA in clone HTSGS30 (ATCC Deposit No. 209232); (e) the complete amino acid sequence excepting the N-terminal methionine encoded by the human cDNA in clone (ATCC Deposit No. 209232); and (f) the complete amino acid sequence of the predicted mature IL-20 polypeptide encoded by the human cDNA in clone (ATCC Deposit No. 209232). The polypeptides of the present invention also include polypeptides having an amino acid sequence at least 80% identical, more preferably at least 90% identical, and still more preferably 95%, 96%, 97%, 98% or 99% identical to those described in (a), (b), (c), (d), (e) or (f), above, as well as polypeptides having an amino acid sequence with at least 90% similarity, and more preferably at least 95% similarity, to those above.

Further polypeptides of the present invention include polypeptides which have at least 90% similarity, more preferably at least 95% similarity, and still more preferably at least 96%, 97%, 98% or 99% similarity to those described above. The polypeptides of the invention also comprise those which are at least 80% identical, more preferably at least 90% or 95% identical, still more preferably at least 96%, 97%, 98% or 99% identical to the polypeptide encoded by the deposited cDNA or to the polypeptide of SEQ ID NO:2, and also include portions of such polypeptides with at least 30 amino acids and more preferably at least 50 amino acids.

A further embodiment of the invention relates to a peptide or polypeptide which comprises the amino acid sequence of a IL-20 polypeptide having an amino acid sequence which contains at least one conservative amino acid substitution, but not more than 50 conservative amino acid substitutions, even more preferably, not more than 40 conservative amino acid substitutions, still more preferably not more than 30 conservative amino acid substitutions, and still even more preferably not more than 20 conservative amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a peptide or polypeptide to have an amino acid sequence which comprises the amino acid sequence of a IL-20 polypeptide, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 conservative amino acid substitutions.

By "% similarity" for two polypeptides is intended a similarity score produced by comparing the amino acid sequences of the two polypeptides using the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711) and the default settings for determining similarity. Bestfit uses the local homology algorithm of Smith and Waterman (Advances in Applied Mathematics 2:482-489, 1981) to find the best segment of similarity between two sequences.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a reference amino acid sequence of an IL-20 polypeptide is intended that the amino acid sequence of the polypeptide is identical to the query sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of the IL-20 polypeptide. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the query sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence shown in SEQ ID NO:2 or to the amino acid sequence encoded by deposited cDNA clone can be determined conventionally

using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed. A preferred method for determing the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is becuase the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be

90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequnce are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

The polypeptide of the present invention could be used as a molecular weight marker on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art.

As described in detail below, the polypeptides of the present invention can also be used to raise polyclonal and monoclonal antibodies, which are useful in assays for detecting IL-20 protein expression as described below or as agonists and antagonists capable of enhancing or inhibiting IL-20 protein function. Further, such polypeptides can be used in the yeast two-hybrid system to "capture" IL-20 protein binding proteins which are also candidate agonists and antagonists according to the present invention. The yeast two hybrid system is described by Fields and Song (*Nature* **340:**245-246 (1989)).

Epitope-Bearing Portions

In another aspect, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes (see, for instance, Geysen, *et al., Proc. Natl. Acad. Sci. USA* **81:**3998-4002 (1983)).

As to the selection of peptides or polypeptides bearing an antigenic epitope (i.e., that contain a region of a protein molecule to which an antibody can bind), it is well known in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein (see, for instance, Sutcliffe, J. G., *et al., Science* **219:**660-666 (1983)). Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins (i.e., immunogenic epitopes) nor to the amino or carboxyl terminals. Antigenic epitope-bearing peptides and polypeptides of the invention are therefore useful to raise antibodies, including monoclonal antibodies, that bind specifically to a polypeptide of the invention (see, for instance, Wilson, *et al., Cell* **37:**767-778 (1984)).

Antigenic epitope-bearing peptides and polypeptides of the invention preferably contain a sequence of at least seven, more preferably at least nine and most preferably between about 15 to about 30 amino acids contained within the amino acid sequence of a polypeptide of the invention. Non-limiting examples of antigenic polypeptides or peptides that can be used to generate IL-20-specific antibodies include: a polypeptide comprising amino acid residues from about Gln-21 to about Arg-29 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Gln-21 to about Gly-41 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Ser-24 to about Gln-32 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Arg-29 to about Pro-37 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Arg-52 to about Glu-60 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Arg-52 to about Met-69 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Glu-61 to about Met-69 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Asn-75 to about Val-86 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Ser-93 to about Trp-101 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Ile-105 to about Pro-113 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Met-132 to about Ser-140 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Arg-150 to about Pro-158 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Pro-156 to about Arg-164 in SEQ ID NO:2, a polypeptide comprising amino acid residues from about Gly-161 to about Met-169 in SEQ ID NO:2, and a polypeptide comprising amino acid residues from about Val-149 to about Ala-167 in SEQ ID NO:2. These polypeptide fragments have been determined to bear antigenic epitopes of the IL-20 protein by the analysis of the Jameson-Wolf antigenic index, as shown in Figure 3, above.

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means (see, for example, Houghten, R. A., *et al., Proc. Natl. Acad. Sci. USA* 82:5131-5135 (1985); and U.S. Patent No. 4,631,211 to Houghten, *et al.* (1986)).

Epitope-bearing peptides and polypeptides of the invention are used to induce antibodies according to methods well known in the art (see, for instance, Sutcliffe, *et al., supra*; Wilson, *et al., supra*; Chow, M., *et al., Proc. Natl. Acad. Sci. USA* **82:**910-914; and Bittle, F. J., *et al., J. Gen. Virol.* **66:**2347-2354 (1985)). Immunogenic epitope-bearing peptides of the invention, i.e., those parts of a protein that elicit an antibody response when the whole protein is the immunogen, are identified according to methods known in the art (see, for instance, Geysen, *et al., supra*). Further still, U.S. Patent No. 5,194,392, issued to Geysen, describes a general method of detecting or determining the sequence of monomers (amino acids or other compounds) which is a topological equivalent of the epitope (i.e., a "mimotope") which is complementary to a particular paratope (antigen binding site) of an antibody of interest. More generally, U.S. Patent No. 4,433,092, issued to Geysen, describes a method of detecting or determining a sequence of monomers which is a topographical equivalent of a ligand which is complementary to the ligand binding site of a particular receptor of interest. Similarly, U.S. Patent No. 5,480,971, issued to

Houghten and colleagues, on Peralkylated Oligopeptide Mixtures discloses linear C1-C7-alkyl peralkylated oligopeptides and sets and libraries of such peptides, as well as methods for using such oligopeptide sets and libraries for determining the sequence of a peralkylated oligopeptide that preferentially binds to an acceptor molecule of interest. Thus, non-peptide analogs of the epitope-bearing peptides of the invention also can be made routinely by these methods.

Fusion Proteins

As one of skill in the art will appreciate, IL-20 polypeptides of the present invention and the epitope-bearing fragments thereof described above can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life *in vivo*. This has been shown, e.g., for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins (EP A 394,827; Traunecker, *et al., Nature* **331:**84-86 (1988)). Fusion proteins that have a disulfide-linked dimeric structure due to the IgG part can also be more efficient in binding and neutralizing other molecules than the monomeric IL-20 protein or protein fragment alone (Fountoulakis, *et al., J. Biochem.* **270:**3958-3964 (1995)).

Antibodies

IL-20 protein-specific antibodies for use in the present invention can be raised against the intact IL-20 protein or an antigenic polypeptide fragment thereof, which may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier.

As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to IL-20 protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding of an intact antibody (Wahl, *et al., J. Nucl. Med.* **24:**316-325 (1983)). Thus, these fragments are preferred.

The antibodies of the present invention may be prepared by any of a variety of methods. For example, cells expressing the IL-20 protein or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of IL-20 protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or IL-20 protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology (Kohler, *et al., Nature* **256:**495 (1975); Kohler, *et al., Eur. J. Immunol.* **6:**511 (1976); Kohler, *et al., Eur. J. Immunol.* **6:**292 (1976); Hammerling, *et*

al., in: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., (1981) pp. 563-681)). In general, such procedures involve immunizing an animal (preferably a mouse) with a IL-20 protein antigen or, more preferably, with a IL-20 protein-expressing cell. Suitable cells can be recognized by their capacity to bind anti-IL-20 protein antibody. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56° C), and supplemented with about 10 µg/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin. The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the American Type Culture Collection, Rockville, Maryland. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands and colleagues (*Gastroenterology* **80:**225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the IL-20 protein antigen.

Alternatively, additional antibodies capable of binding to the IL-20 protein antigen may be produced in a two-step procedure through the use of anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and that, therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, IL-20 protein-specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the IL-20 protein-specific antibody can be blocked by the IL-20 protein antigen. Such antibodies comprise anti-idiotypic antibodies to the IL-20 protein-specific antibody and can be used to immunize an animal to induce formation of further IL-20 protein-specific antibodies.

It will be appreciated that Fab and F(ab')2 and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). Alternatively, IL-20 protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For *in vivo* use of anti-IL-20 in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art (Morrison, *Science* **229:**1202 (1985); Oi, *et al., BioTechniques* **4:**214 (1986); Cabilly, *et al.*, U.S. Patent No. 4,816,567; Taniguchi, *et al.*, EP 171496; Morrison, *et al.*, EP 173494; Neuberger, *et al.*, WO 8601533; Robinson, *et al.*, WO 8702671; Boulianne, *et al., Nature* **312:**643 (1984); Neuberger, *et al., Nature* **314:**268 (1985).

Immune System-Related Disorders

Diagnosis

The present inventors have discovered that IL-20 is expressed in the thymus. For a number of immune system-related disorders, substantially altered (increased or decreased) levels of IL-20 gene expression can be detected in immune system tissue or other cells or bodily fluids (e.g., sera, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to a "standard" IL-20 gene expression level, that is, the IL-20 expression level in immune system tissues or bodily fluids from an individual not having the immune system disorder. Thus, the invention provides a diagnostic method useful during diagnosis of a immune system disorder, which involves measuring the expression level of the gene encoding the IL-20 protein in immune system tissue or other cells or body fluid from an individual and comparing the measured gene expression level with a standard IL-20 gene expression level, whereby an increase or decrease in the gene expression level compared to the standard is indicative of an immune system disorder.

In particular, it is believed that certain tissues in mammals with cancer of the immune system express significantly enhanced levels of the IL-20 protein and mRNA encoding the IL-20 protein when compared to a corresponding "standard" level. Further, it is believed that enhanced levels of the IL-20 protein can be detected in certain body fluids (e.g., sera, plasma, urine, and spinal fluid) from mammals with such a cancer when compared to sera from mammals of the same species not having the cancer.

Thus, the invention provides a diagnostic method useful during diagnosis of a immune system disorder, including cancers of this system, which involves measuring the expression level of the gene encoding the IL-20 protein in immune system tissue or other cells or body fluid from an individual and comparing the measured gene expression level with a standard IL-20 gene expression level, whereby an increase or decrease in the gene expression level compared to the standard is indicative of an immune system disorder.

Where a diagnosis of a disorder in the immune system, including diagnosis of a tumor, has already been made according to conventional methods, the present invention is useful as a prognostic indicator, whereby patients exhibiting enhanced IL-20 gene expression will experience a worse clinical outcome relative to patients expressing the gene at a level nearer the standard level.

By "assaying the expression level of the gene encoding the IL-20 protein" is intended qualitatively or quantitatively measuring or estimating the level of the IL-20 protein or the level of the mRNA encoding the IL-20 protein in a first biological sample either directly (e.g., by determining or estimating absolute protein level or mRNA level) or relatively (e.g., by comparing to the IL-20 protein level or mRNA level in a second biological sample). Preferably, the IL-20 protein level or mRNA level in the first biological sample is measured or estimated and compared to a standard IL-20 protein level or mRNA level, the standard being taken from a second

biological sample obtained from an individual not having the disorder or being determined by averaging levels from a population of individuals not having a disorder of the immune system. As will be appreciated in the art, once a standard IL-20 protein level or mRNA level is known, it can be used repeatedly as a standard for comparison.

By "biological sample" is intended any biological sample obtained from an individual, body fluid, cell line, tissue culture, or other source which contains IL-20 protein or mRNA. As indicated, biological samples include body fluids (such as sera, plasma, urine, synovial fluid and spinal fluid) which contain free mature IL-20 protein, immune system tissue, and other tissue sources found to express complete or mature IL-20 or an IL-20 receptor. Methods for obtaining tissue biopsies and body fluids from mammals are well known in the art. Where the biological sample is to include mRNA, a tissue biopsy is the preferred source.

The present invention is useful for diagnosis or treatment of various immune system-related disorders in mammals, preferably humans. Such disorders include tumors, cancers, interstitial lung disease (such as Langerhans cell granulomatosis), and any disregulation of immune cell function including, but not limited to, autoimmunity, arthritis, leukemias, lymphomas, immunosuppression, immunity, humoral immunity, inflammatory bowel disease, myelo suppression, and the like.

Total cellular RNA can be isolated from a biological sample using any suitable technique such as the single-step guanidinium-thiocyanate-phenol-chloroform method described by Chomczynski and Sacchi (*Anal. Biochem.* **162:**156-159 (1987)). Levels of mRNA encoding the IL-20 protein are then assayed using any appropriate method. These include Northern blot analysis, S1 nuclease mapping, the polymerase chain reaction (PCR), reverse transcription in combination with the polymerase chain reaction (RT-PCR), and reverse transcription in combination with the ligase chain reaction (RT-LCR).

Assaying IL-20 protein levels in a biological sample can occur using antibody-based techniques. For example, IL-20 protein expression in tissues can be studied with classical immunohistological methods (Jalkanen, M., *et al., J. Cell. Biol.* **101:**976-985 (1985); Jalkanen, M., *et al., J. Cell. Biol.* **105:**3087-3096 (1987)). Other antibody-based methods useful for detecting IL-20 protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (^{125}I, ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^{3}H), indium (^{112}In), and technetium (^{99m}Tc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying IL-20 protein levels in a biological sample obtained from an individual, IL-20 protein can also be detected *in vivo* by imaging. Antibody labels or markers for *in vivo* imaging of IL-20 protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and

ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

A IL-20 protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ^{131}I, ^{112}In, ^{99m}Tc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously or intraperitoneally) into the mammal to be examined for immune system disorder. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of ^{99m}Tc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain IL-20 protein. *In vivo* tumor imaging is described by Burchiel and coworkers (Chapter 13 in *Tumor Imaging: The Radiochemical Detection of Cancer*, Burchiel, S. W. and Rhodes, B. A., eds., Masson Publishing Inc. (1982)).

Treatment

As noted above, IL-20 polynucleotides and polypeptides are useful for diagnosis of conditions involving abnormally high or low expression of IL-20 activities. Given the cells and tissues where IL-20 is expressed as well as the activities modulated by IL-20, it is readily apparent that a substantially altered (increased or decreased) level of expression of IL-20 in an individual compared to the standard or "normal" level produces pathological conditions related to the bodily system(s) in which IL-20 is expressed and/or is active.

Based on its ability to alter expression of the cellular transcription factor NF-kB and induce IL-6 expression, IL-20 may be used to treat B-cell neoplasms, including chronic lymphocytic leukemia (CLL) and B-lymphocytic leukemia (BLL). Further, an IL-20-mediated induction of IL-6 expression can also be used to activate mature lymphoid cells, which have cytolytic activities. As a result, IL-20 can be used as an anticancer and antiviral treatment. Various immunodeficiencies, for example, in T- and B-lymphocytes, or immune disorders, for example, rheumatoid arthritis, may also be beneficially effected by treatment with IL-20. Immunodeficiencies such as leukopenia, a reduction in the number of circulating leukocytes in the peripheral blood, may be the result of viral infections, for example, HIV, severe exposure to radiation, side effects of cancer therapy or the result of other medical treatment. Therapeutic treatment of leukopenia with IL-20 compositionsmay avoid undesirable side effects caused by treatment with presently available drugs. Other conditions susceptible for IL-20 include patients recovering from bone marrow transplants. IL-20 may also be used to augment the humoral or cellular immune response *in vivo* in coadministration with other therapeutic agents. For example, IL-20 may be used to enhance the efficacy of viral antigen vaccines, such as HIV or tumor antigen vaccines.

Primarily through its effects on expression of the cellular transcription factor NF-kB and IL-6 expression, IL-20 also functions as a hybridomal growth factor in culture medium for hybridoma cell lines to increase the yield thereof.

IL-20 may also be useful in immunotherapeutic and anti-inflammation compositions. IL-20 may also be used for the treatment of patients suffering from chemotherapy from bone marrow transplants. IL-20 may be further used to treat corneal damage, keratitis, and ulcers.

Among the other treatments IL-20 may be used for include conditions such as thrombocytopenia, in which IL-20 will enhance differentiation into platelet producing cells. IL-20 may also be used to restore neutrophil and platelet counts in treatment of cancer and in bone marrow transplantation.

IL-20 may also be used to induce liver cells to produce a number of proteins called "acute phase proteins". The acute phase proteins are usually induced after an acute insult, such as traumatic or bacterial shock. Accordingly, administration of IL-20 may be beneficial in promoting recovery.

IL-20 may also be used in cell transplant therapy including autogenous bone marrow graft therapy.

IL-20 may also be employed to enhance erythropoietin production for treating anemias associated with inflammation, renal failure, AIDS, and cancer.

IL-20 may be used, alone or incombination with other therapeutic products, in the treatment of diseases characterized by a decreased level of either myeloid or lymphoid cells of the hematopoietic system. This protein may also be capable of stimulating accessory and mature cells, for example, monocytes to produce other hematopoietic-like factors which, in turn stimulate the formation of colonies of other hematopoietic cells, as well as other hematopoietic-like activities.

It is well-known in the art that, in addition to a specific cellular function, cellular receptor molecules may also often be exploited by a virus as a means of initiating entry into a potential host cell. For example, it was recently discovered by Wu and colleagues (*J. Exp. Med.* **185:**1681-1691 (1997)) that the cellular chemokine receptor CCR5 functions not only as a cellular chemokine receptor, but also as a receptor for macrophage-tropic human immunodeficiency virus (HIV)-1. In addition, RANTES, MIP-1a, and MIP-1b, which are agonists for the cellular chemokine receptor CCR5, inhibit entry of various strains of HIV-1 into susceptible cell lines (Cocchi, F., *et al., Science* **270:**1811-1815 (1995)). Thus, the invention also provides a method of treating an individual exposed to, or infected with, a virus through the prophylactic or therapeutic administration of IL-20, or an agonist or antagonist thereof, to block or disrupt the interaction of a viral particle with the IL-20 receptor and, as a result, block the initiation or continuation of viral infectivity.

The IL-20 of the present invention binds to the IL-20 receptor and, as such, is likely to block immuno-tropic viral infections. Further, expression patterns of cytokines and cytokine receptors suggest that the IL-20 receptor is expressed primarily in hematopoietic and neural

tissues. These observations further suggest that agonists and antagonists, including ligands, of IL-20 may be useful as a method of blocking or otherwise regulating the infectivity of immunotropic and neurotropic viral infections. A non-limiting list of viruses which infect humans and can infect cells of the hematopoietic and nervous systems includes such retroviruses as HIV-1, HIV-2, human T-cell lymphotropic virus (HTLV)-I, and HTLV-II, as well as other DNA and RNA viruses such as herpes simplex virus (HSV)-1, HSV-2, HSV-6, cytomegalovirus (CMV), Epstein-Barr virus (EBV), herpes samirii, adenoviruses, rhinoviruses, influenza viruses, reoviruses, and the like.

The ability of the IL-20 of the present invention, or agonists or antagonists thereof, to prophylactically or therapeutically block viral infection may be easily tested by the skilled artisan. For example, Simmons and coworkers (Science 276:276-279 (1997)) and Arenzana-Seisdedos and colleagues (*Nature* **383**:400 (1996)) each outline a method of observing suppression of HIV-1 infection by an antagonist of the CCR5 chemokine receptor and of the CC chemokine RANTES, respectively, in cultured peripheral blood mononuclear cells. Cells are cultured and infected with a virus, HIV-1 in both cases noted above. An agonist or antagonist of the CC chemokine or its receptor is then immediately added to the culture medium. Evidence of the ability of the agonist or antagonist of the chemokine or cellular receptor is determined by evaluating the relative success of viral infection at 3, 6, and 9 days postinfection.

Administration of a pharmaceutical composition comprising an amount of an isolated IL-20, or an agonist or antagonist thereof, of the invention to an individual either infected with a virus or at risk for infection with a virus is performed as described below.

It will also be appreciated by one of ordinary skill that, since the IL-20 protein of the invention is a member of the cytokine family of polypeptides, the mature secreted form of the protein may be released in soluble form from the cells which express the IL-20 by proteolytic cleavage. Therefore, when IL-20 mature form is added from an exogenous source to cells, tissues or the body of an individual, the protein will exert its physiological activities on its target cells of that individual.

Therefore, it will be appreciated that conditions caused by a decrease in the standard or normal level of IL-20 activity in an individual, particularly disorders of the immune system, can be treated by administration of IL-20 polypeptide (in the form of the mature protein). Thus, the invention also provides a method of treatment of an individual in need of an increased level of IL-20 activity comprising administering to such an individual a pharmaceutical composition comprising an amount of an isolated IL-20 polypeptide of the invention, particularly a mature form of the IL-20 protein of the invention, effective to increase the IL-20 activity level in such an individual.

Since IL-20 is a novel homologue of the recently described cytokine IL-17, it will have a wide range of cytokine-like activities. IL-20 may be employed to enhance host defenses against resistant chronic and acute infections, for example, mycobacterial infections via the attraction and activation of microbicidal leukocytes. IL-20 may also be employed to increase T-cell

proliferation by the stimulation of IL-2 biosynthesis for the treatment of T-cell mediated auto-immune diseases and lymphocytic leukemias. IL-20 may also be employed to regulate hematopoiesis, by regulating the activation and differentiation of various hematopoietic progenitor cells, for example, to release mature leukocytes from the bone marrow following chemotherapy, i.e., in stem cell mobilization. IL-20 may also be employed to treat sepsis. Also, stimulation of the cell by IL-20 strongly induces IL-6 expression. IL-6 is a potent growth factor for myelomas, plasmacytomas, and hybridomas and is involved in the growth of Lennert's Lymphoma T-cells. As a result, IL-20 and IL-20 agonists may be used in the treatment of such cancers, analogous disease states, and others known to those of skill in the art.

Schwann cells, and microglia and astrocytes are the immunocompetent cells of the peripheral and central nervous systems, respectively, that secrete a variety of immune and inflammatory mediators. Inflammatory processes involving reactive microglia, e.g. those associated with the lesions found following stroke or in multiple sclerosis, and with beta-amyloid containing plaques in Alzheimer's Disease, have been proposed to contribute to the neuronal pathology characteristic of these clinical conditions. In the peripheral nervous system, there is increasing evidence that Schwann cells play an essential role in the pathogenesis associated with autoimmune inflammatory peripheral nerve disease as well as other demylinating diseases such as Guillain-Barr syndrome. Furthermore, all three cell types are targets for numerous interleukins, including IL-20.

Formulations

The IL-20 polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with IL-20 polypeptide alone), the site of delivery of the IL-20 polypeptide composition, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" of IL-20 polypeptide for purposes herein is thus determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of IL-20 polypeptide administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the IL-20 polypeptide is typically administered at a dose rate of about 1 µg/kg/hour to about 50 µg/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

Pharmaceutical compositions containing the IL-20 of the invention may be administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by

powders, ointments, drops or transdermal patch), bucally, or as an oral or nasal spray. By "pharmaceutically acceptable carrier" is meant a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

The IL-20 polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or mirocapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman, U., *et al., Biopolymers* **22:**547-556 (1983)), poly (2-hydroxyethyl methacrylate; Langer, R., *et al., J. Biomed. Mater. Res.* **15:**167-277 (1981), and Langer, R., *Chem. Tech.* **12:**98-105 (1982)), ethylene vinyl acetate (Langer, R., *et al., Id.*) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release IL-20 polypeptide compositions also include liposomally entrapped IL-20 polypeptide. Liposomes containing IL-20 polypeptide are prepared by methods known in the art (DE 3,218,121; Epstein, *et al., Proc. Natl. Acad. Sci. (USA)* **82:**3688-3692 (1985); Hwang, *et al., Proc. Natl. Acad. Sci. (USA)* **77:**4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324). Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal IL-20 polypeptide therapy.

For parenteral administration, in one embodiment, the IL-20 polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

Generally, the formulations are prepared by contacting the IL-20 polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as

polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The IL-20 polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of IL-20 polypeptide salts.

IL-20 polypeptide to be used for therapeutic administration must be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic IL-20 polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

IL-20 polypeptide ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous IL-20 polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized IL-20 polypeptide using bacteriostatic water-for-injection (WFI).

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Agonists and Antagonists - Assays and Molecules

The invention also provides a method of screening compounds to identify those which enhance or block the action of IL-20 on cells, such as its interaction with IL-20-binding molecules such as receptor molecules. An agonist is a compound which increases the natural biological functions of IL-20 or which functions in a manner similar to IL-20, while antagonists decrease or eliminate such functions.

In another aspect of this embodiment the invention provides a method for identifying a receptor protein or other ligand-binding protein which binds specifically to an IL-20 polypeptide. For example, a cellular compartment, such as a membrane or a preparation thereof, may be prepared from a cell that expresses a molecule that binds IL-20. The preparation is incubated with labeled IL-20 and complexes of IL-20 bound to the receptor or other binding protein are

isolated and characterized according to routine methods known in the art. Alternatively, the IL-20 polypeptide may be bound to a solid support so that binding molecules solubilized from cells are bound to the column and then eluted and characterized according to routine methods.

In the assay of the invention for agonists or antagonists, a cellular compartment, such as a membrane or a preparation thereof, may be prepared from a cell that expresses a molecule that binds IL-20, such as a molecule of a signaling or regulatory pathway modulated by IL-20. The preparation is incubated with labeled IL-20 in the absence or the presence of a candidate molecule which may be a IL-20 agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labeled ligand. Molecules which bind gratuitously, i.e., without inducing the effects of IL-20 on binding the IL-20 binding molecule, are most likely to be good antagonists. Molecules that bind well and elicit effects that are the same as or closely related to IL-20 are agonists.

IL-20-like effects of potential agonists and antagonists may by measured, for instance, by determining activity of a second messenger system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of IL-20 or molecules that elicit the same effects as IL-20. Second messenger systems that may be useful in this regard include but are not limited to AMP guanylate cyclase, ion channel or phosphoinositide hydrolysis second messenger systems.

Another example of an assay for IL-20 antagonists is a competitive assay that combines IL-20 and a potential antagonist with membrane-bound IL-20 receptor molecules or recombinant IL-20 receptor molecules under appropriate conditions for a competitive inhibition assay. IL-20 can be labeled, such as by radioactivity, such that the number of IL-20 molecules bound to a receptor molecule can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a receptor molecule, without inducing IL-20-induced activities, thereby preventing the action of IL-20 by excluding IL-20 from binding.

Other potential antagonists include antisense molecules. Antisense technology can be used to control gene expression through antisense DNA or RNA or through triple-helix formation. Antisense techniques are discussed in a number of studies (for example, Okano, *J. Neurochem.* **56:**560 (1991); "Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression." CRC Press, Boca Raton, FL (1988)). Triple helix formation is discussed in a number of studies, as well (for instance, Lee, *et al., Nucleic Acids Research* **6:**3073 (1979); Cooney, *et al., Science* **241:**456 (1988); Dervan, *et al., Science* **251:**1360 (1991)). The methods are based on binding of a polynucleotide to a complementary DNA or RNA. For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the

present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of IL-20. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into IL-20 polypeptide. The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of IL-20 protein.

The agonists and antagonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as described above.

The antagonists may be employed for instance to inhibit the activation of macrophages and their precursors, and of neutrophils, basophils, B lymphocytes and some T-cell subsets, e.g., activated and CD8 cytotoxic T cells and natural killer cells, in certain auto-immune and chronic inflammatory and infective diseases. Examples of auto-immune diseases include multiple sclerosis, and insulin-dependent diabetes. The antagonists may also be employed to treat infectious diseases including silicosis, sarcoidosis, idiopathic pulmonary fibrosis by preventing the activation of mononuclear phagocytes. They may also be employed to treat idiopathic hyper-eosinophilic syndrome by preventing eosinophil production. Antagonists may also be employed to treat rheumatoid arthritis by preventing the activation of monocytes in the synovial fluid in the joints of patients. Monocyte activation plays a significant role in the pathogenesis of both degenerative and inflammatory arthropathies. The antagonists may be employed to interfere with the deleterious cascades attributed primarily to IL-1 and TNF, which prevents the biosynthesis of other inflammatory cytokines. In this way, the antagonists may be employed to prevent inflammation. Antibodies against IL-20 may be employed to bind to and inhibit IL-20 activity to treat such conditions described above. Any of the above antagonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as hereinafter described.

Gene Mapping

The nucleic acid molecules of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

In certain preferred embodiments in this regard, the cDNA herein disclosed is used to clone genomic DNA of a IL-20 protein gene. This can be accomplished using a variety of well known techniques and libraries, which generally are available commercially. The genomic DNA then is used for *in situ* chromosome mapping using well known techniques for this purpose.

In addition, in some cases, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3' untranslated region of the gene is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Fluorescence *in situ* hybridization ("FISH") of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with probes from the cDNA as short as 50 or 60 bp (for a review of this technique, see Verma, *et al., Human Chromosomes: A Manual Of Basic Techniques*, Pergamon Press, New York (1988)).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, on the World Wide Web (McKusick, V. *Mendelian Inheritance In Man*, available on-line through Johns Hopkins University, Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Examples

Example 1(a): Expression and Purification of "His-tagged" IL-20 in E. coli

The bacterial expression vector pQE9 (pD10) is used for bacterial expression in this example. (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311). pQE9 encodes ampicillin antibiotic resistance ("Ampr") and contains a bacterial origin of replication ("ori"), an IPTG inducible promoter, a ribosome binding site ("RBS"), six codons encoding histidine residues that allow affinity purification using nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin sold by QIAGEN, Inc., *supra,* and suitable single restriction enzyme cleavage sites. These elements are arranged such that an inserted DNA fragment encoding a polypeptide expresses that polypeptide with the six His residues (i.e., a "6 X His tag") covalently linked to the amino terminus of that polypeptide.

The DNA sequence encoding the desired portion of the IL-20 protein comprising the mature form of the IL-20 amino acid sequence is amplified from the deposited cDNA clone using

PCR oligonucleotide primers which anneal to the amino terminal sequences of the desired portion of the IL-20 protein and to sequences in the deposited construct 3' to the cDNA coding sequence. Additional nucleotides containing restriction sites to facilitate cloning in the pQE9 vector are added to the 5' and 3' primer sequences, respectively.

For cloning the mature form of the IL-20 protein, the 5' primer has the sequence 5' GAT CGC GGA TCC CAG CCC AGG AGC CCC AAA AGC AAG AGG AAG-3' (SEQ ID NO:5) containing the underlined *Bam* I restriction site followed by 30 nucleotides of the amino terminal coding sequence of the mature IL-20 sequence in SEQ ID NO:2. One of ordinary skill in the art would appreciate, of course, that the point in the protein coding sequence where the 5' primer begins may be varied to amplify a DNA segment encoding any desired portion of the complete IL-20 protein shorter or longer than the mature form of the protein. The 3' primer has the sequence 5' GAT CGC AAG CTT CAG GTT TAT CAG AAG ATG CAG GTG CAG CCC ACA GC-3' (SEQ ID NO:6) containing the underlined *Hind* III restriction site followed by 35 nucleotides complementary to the 3' end of the coding sequence of the IL-20 DNA sequence in Figure 1.

The amplified IL-20 DNA fragment and the vector pQE9 are digested with *Bam* I and *Hind* III and the digested DNAs are then ligated together. Insertion of the IL-20 DNA into the restricted pQE9 vector places the IL-20 protein coding region downstream from the IPTG-inducible promoter and in-frame with an initiating AUG and the six histidine codons.

The ligation mixture is transformed into competent *E. coli* cells using standard procedures such as those described by Sambrook and colleagues (*Molecular Cloning: a Laboratory Manual, 2nd Ed.*; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989)). *E. coli* strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses the lac repressor and confers kanamycin resistance ("Kanr"), is used in carrying out the illustrative example described herein. This strain, which is only one of many that are suitable for expressing IL-20 protein, is available commercially (QIAGEN, Inc., *supra*). Transformants are identified by their ability to grow on LB plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis, PCR and DNA sequencing.

Clones containing the desired constructs are grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 µg/ml) and kanamycin (25 µg/ml). The O/N culture is used to inoculate a large culture, at a dilution of approximately 1:25 to 1:250. The cells are grown to an optical density at 600 nm ("OD600") of between 0.4 and 0.6. Isopropyl-β-D-thiogalactopyranoside ("IPTG") is then added to a final concentration of 1 mM to induce transcription from the lac repressor sensitive promoter, by inactivating the lacI repressor. Cells subsequently are incubated further for 3 to 4 hours. Cells then are harvested by centrifugation.

The cells are then stirred for 3-4 hours at 4°C in 6M guanidine-HCl, pH 8. The cell debris is removed by centrifugation, and the supernatant containing the IL-20 is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist, 1995, QIAGEN, Inc., *supra*). Briefly the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the IL-20 is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins can be eluted by the addition of 250 mM immidazole. Immidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4° C or frozen at -80° C.

The following alternative method may be used to purify IL-20 expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10°C and the cells are harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

The cells ware then lysed by passing the solution through a microfluidizer (Microfuidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 x *g* for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 x *g* centrifugation for 15 min., the pellet is discarded and the IL-20 polypeptide-containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

Following high speed centrifugation (30,000 x *g*) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

To clarify the refolded IL-20 polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 μm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 mm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the IL-20 polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A_{280} monitoring of the effluent. Fractions containing the IL-20 polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant IL-20 polypeptide exhibits greater than 95% purity after the above refolding and purification steps. No major contaminant bands are observed from Commassie blue stained 16% SDS-PAGE gel when 5 μg of purified protein is loaded. The purified protein is also tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

IL-20 polypeptide was expressed in the W3110 *E. coli* strain and was found in inclusion bodies. The IL-20 present in inclusion bodies was solubilized with 3-4 M guanidine in 0.1 M Na phosphate buffer, pH 8, 10 mM EDTA. Extraction with 4 M or 8 M urea resulted in apparent proteolytic degradation of the induced IL-20 protein band upon dilution of the urea concentration to 4M or after dialysis against 50 mM sodium acetate buffer, pH 6, 0.1M NaCl, 2 mM EDTA . The 4 M guanidine extract of IL-20 inclusion bodies was found to retain solubility and remain intact after dilution to 0.3 M guanidine if extracted overnight in the presence of 10 mM DTT or 5 mM cysteine.

As IL-20 contains 8 cysteine residues, it was of interest to analyze whether or not the reason for the protein being insoluble after removal of guanidine was due to the presence of

disulfide-linked aggregates. SDS-PAGE analysis of the 0.3 M guanidine solubilized fraction under non-reducing conditions resulted in a higher electrophorectic mobility versus reduced sample suggestive of intramolecular disulfide bond formation, in addition, no high molecular weight species were noted. Also, size exclusion analysis in the presence of 0.4 M guanidine indicated that the protein is either a monomer or dimer and is not present as a high molecular weight homo- or hetero- protein aggregates. The only time IL-20 disulfide-linked aggregates were found was when inclusion bodies were solubilized in the presence of reduced/oxidized glutathione.

Lowering of the guanidine concentration by dialysis against 0.1 M NaCl in buffers at pH 5.5, 6, 8, or 9 all resulted in precipitation of the protein along with other impurities. However, after dialysis against pH 3.5 acetate buffer in 0.125 M NaCl there was a significant amount of soluble IL-20 which was about ~70 % pure.

Example 2: Cloning and Expression of IL-20 protein in a Baculovirus Expression System

In this illustrative example, the plasmid shuttle vector pA2 is used to insert the cloned DNA encoding complete protein, including its naturally associated secretory signal (leader) sequence, into a baculovirus to express the mature IL-20 protein, using standard methods as described by Summers and colleagues (*A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas Agricultural Experimental Station Bulletin No. 1555 (1987)). This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as *Bam* HI, *Xba* I and *Asp* 718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

Many other baculovirus vectors could be used in place of the vector above, such as pAc373, pVL941 and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, by Luckow and coworkers (*Virology* **170:**31-39 (1989)).

The cDNA sequence encoding the full length IL-20 protein in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence shown in SEQ ID NO:2, is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene. The 5' primer has the sequence 5' GAT CGC GGA TCC GCC ATC

ATG GAC TGG CCT CAC AAC CTG CTG TTT CTT CTT AC 3' (SEQ ID NO:7) containing the underlined *Bam* HI restriction enzyme site, an efficient signal for initiation of translation in eukaryotic cells (Kozak, M., *J. Mol. Biol.* **196:**947-950 (1987)), followed by 35 of nucleotides of the sequence of the complete IL-20 protein shown in Figure 1, beginning with the AUG initiation codon. The 3' primer has the sequence 5' GAT CGC GGT ACC CAG GTT TAT CAG AAG ATG CAG GTG CAG CCC ACA GC 3' (SEQ ID NO:8) containing the underlined *Asp* 718 restriction site followed by 35 nucleotides complementary to the 3' noncoding sequence in Figure 1.

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with *Bam* HI and *Asp* 718 and again is purified on a 1% agarose gel. This fragment is designated herein F1.

The plasmid is digested with the restriction enzymes *Bam* HI and *Asp* 718 and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.). This vector DNA is designated herein "V1".

Fragment F1 and the dephosphorylated plasmid V1 are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Statagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria are identified that contain the plasmid with the human IL-20 gene by digesting DNA from individual colonies using *Bam* HI and *Asp* 718 and then analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing. This plasmid is designated herein pA2IL-20.

Five µg of the plasmid pA2IL-20 is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner and colleaguew (*Proc. Natl. Acad. Sci. USA* **84:**7413-7417 (1987)). One µg of BaculoGold™ virus DNA and 5 µg of the plasmid pA2IL-20 are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27°C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27°C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith (*supra*). An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which

produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10). After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4°C. The recombinant virus is called V-IL-20.

To verify the expression of the IL-20 gene Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-IL-20 at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 µCi of ^{35}S-methionine and 5 µCi ^{35}S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the mature form of the IL-20 protein, and thus the cleavage point and length of the naturally associated secretory signal peptide.

Example 3: Cloning and Expression of IL-20 in Mammalian Cells

A typical mammalian expression vector contains the promoter element, which mediates the initiation of transcription of mRNA, the protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription can be achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter). Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) and pBC12MI (ATCC 67109). Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the gene can be expressed in stable cell lines that contain the gene integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful to develop cell lines that carry several hundred or even several thousand copies of the gene of interest. Another useful selection marker is the enzyme glutamine synthase (GS; Murphy, *et al., Biochem J.* **227:**277-279 (1991); Bebbington, *et al., Bio/Technology* **10:**169-175 (1992)). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

The expression vectors pC1 and pC4 contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen, *et al., Mol. Cel. Biol.* **5:**438-447 (1985)) plus a fragment of the CMV-enhancer (Boshart, *et al., Cell* **41:**521-530 (1985)). Multiple cloning sites, e.g., with the restriction enzyme cleavage sites *Bam* HI, *Xba* I and *Asp* 718, facilitate the cloning of the gene of interest. The vectors contain in addition the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene.

Example 3(a): Cloning and Expression in COS Cells

The expression plasmid, pIL-20HA, is made by cloning a portion of the cDNA encoding the mature form of the IL-20 protein into the expression vector pcDNAI/Amp or pcDNAIII (which can be obtained from Invitrogen, Inc.).

The expression vector pcDNAI/amp contains: (1) an *E. coli* origin of replication effective for propagation in *E. coli* and other prokaryotic cells; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin of replication for propagation in eukaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron; (5) several codons encoding a hemagglutinin fragment (i.e., an "HA" tag to facilitate purification) followed by a termination codon and polyadenylation signal arranged so that a cDNA can be conveniently placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein described by Wilson and colleagues (*Cell* **37:**767 (1984)). The fusion of the HA tag to the target protein allows easy detection and recovery of the recombinant protein with an antibody that recognizes the HA epitope. pcDNAIII contains, in addition, the selectable neomycin marker.

A DNA fragment encoding the complete IL-20 polypeptide is cloned into the polylinker region of the vector so that recombinant protein expression is directed by the CMV promoter.

The plasmid construction strategy is as follows. The IL-20 cDNA of the deposited clone is amplified using primers that contain convenient restriction sites, much as described above for construction of vectors for expression of IL-20 in *E. coli*. Suitable primers include the following, which are used in this example. The 5' primer, containing the underlined *Asp* 718 site, a Kozak sequence, an AUG start codon, and 35 nucleotides of the 5' coding region of the complete IL-20 polypeptide, has the following sequence: 5' GAT CGC GGT ACC GCC ATC ATG GAC TGG CCT CAC AAC CTG CTG TTT CTT CTT AC 3' (SEQ ID NO:9). The 3' primer, containing the underlined *Bam* HI and 35 of nucleotides complementary to the 3' coding sequence immediately before the stop codon, has the following sequence: 5' GAT CGC GGA TCC CAG GTT TAT CAG AAG ATG CAG GTG CAG CCC ACA GC 3' (SEQ ID NO:10).

The PCR amplified DNA fragment and the vector, pcDNAI/Amp, are digested with *Bam* HI and *Asp* 718 and then ligated. The ligation mixture is transformed into *E. coli* strain SURE (Stratagene Cloning Systems, La Jolla, CA 92037), and the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis or other means for the presence of the fragment encoding the complete IL-20 polypeptide

For expression of recombinant IL-20, COS cells are transfected with an expression vector, as described above, using DEAE-dextran, as described, for instance, by Sambrook and coworkers (*Molecular Cloning: a Laboratory Manual*, Cold Spring Laboratory Press, Cold Spring Harbor, New York (1989)). Cells are incubated under conditions for expression of IL-20 by the vector.

Expression of the IL-20-HA fusion protein is detected by radiolabeling and immunoprecipitation, using methods described in, for example Harlow and colleagues (*Antibodies: A Laboratory Manual, 2nd Ed.*; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988)). To this end, two days after transfection, the cells are labeled by incubation in media containing ^{35}S-cysteine for 8 hours. The cells and the media are collected, and the cells are washed and the lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson and colleagues (*supra*). Proteins are precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The precipitated proteins then are analyzed by SDS-PAGE and autoradiography. An expression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

Example 3(b): Cloning and Expression in CHO Cells

The vector pC4 is used for the expression of IL-20 polypeptide. Plasmid pC4 is a derivative of the plasmid pSV2-dhfr (ATCC Accession No. 37146). The plasmid contains the

mouse DHFR gene under control of the SV40 early promoter. Chinese hamster ovary- or other cells lacking dihydrofolate activity that are transfected with these plasmids can be selected by growing the cells in a selective medium (alpha minus MEM, Life Technologies) supplemented with the chemotherapeutic agent methotrexate. The amplification of the DHFR genes in cells resistant to methotrexate (MTX) has been well documented (see, e.g., Alt, F. W., *et al., J. Biol. Chem.* **253:**1357-1370 (1978); Hamlin, J. L. and Ma, C. *Biochem. et Biophys. Acta,* **1097:**107-143 (1990); Page, M. J. and Sydenham, M. A. *Biotechnology* **9:**64-68 (1991)). Cells grown in increasing concentrations of MTX develop resistance to the drug by overproducing the target enzyme, DHFR, as a result of amplification of the DHFR gene. If a second gene is linked to the DHFR gene, it is usually co-amplified and over-expressed. It is known in the art that this approach may be used to develop cell lines carrying more than 1,000 copies of the amplified gene(s). Subsequently, when the methotrexate is withdrawn, cell lines are obtained which contain the amplified gene integrated into one or more chromosome(s) of the host cell.

Plasmid pC4 contains for expressing the gene of interest the strong promoter of the long terminal repeat (LTR) of the Rouse Sarcoma Virus (Cullen, *et al., Mol. Cell. Biol.* **5:**438-447 (1985)) plus a fragment isolated from the enhancer of the immediate early gene of human cytomegalovirus (CMV; Boshart, *et al., Cell* **41:**521-530 (1985)). Downstream of the promoter are the following single restriction enzyme cleavage sites that allow the integration of the genes: *Bam* HI, *Xba* I, and *Asp* 718. Behind these cloning sites the plasmid contains the 3' intron and polyadenylation site of the rat preproinsulin gene. Other high efficiency promoters can also be used for the expression, e.g., the human ß-actin promoter, the SV40 early or late promoters or the long terminal repeats from other retroviruses, e.g., HIV and HTLVI. Clontech's Tet-Off and Tet-On gene expression systems and similar systems can be used to express the IL-20 polypeptide in a regulated way in mammalian cells (Gossen, M., and Bujard, H. *Proc. Natl. Acad. Sci. USA* **89:**5547-5551 (1992)). For the polyadenylation of the mRNA other signals, e.g., from the human growth hormone or globin genes can be used as well. Stable cell lines carrying a gene of interest integrated into the chromosomes can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

The plasmid pC4 is digested with the restriction enzymes *Bam* HI and *Asp* 718 and then dephosphorylated using calf intestinal phosphates by procedures known in the art. The vector is then isolated from a 1% agarose gel.

The DNA sequence encoding the complete IL-20 polypeptide is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the desired portion of the gene. The 5' primer containing the underlined *Bam* HI site, a Kozak sequence, an AUG start codon, and 35 nucleotides of the 5' coding region of the complete IL-20 polypeptide, has the

following sequence: 5' GAT CGC GGA TCC GCC ATC ATG GAC TGG CCT CAC AAC CTG CTG TTT CTT CTT AC 3' (SEQ ID NO:7). The 3' primer, containing the underlined *Asp* 718 and 35 of nucleotides complementary to the 3' coding sequence immediately before the stop codon as shown in Figure 1 (SEQ ID NO:1), has the following sequence: 5' GAT CGC GGT ACC CAG GTT TAT CAG AAG ATG CAG GTG CAG CCC ACA GC 3' (SEQ ID NO:8).

The amplified fragment is digested with the endonucleases *Bam* HI and *Asp* 718 and then purified again on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC4 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene are used for transfection. Five µg of the expression plasmid pC4 is cotransfected with 0.5 µg of the plasmid pSVneo using lipofectin (Felgner, *et al., supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of metothrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 µM, 2 µM, 5 µM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100-200 µM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 4: Tissue distribution of IL-20 mRNA expression

Northern blot analysis is carried out to examine IL-20 gene expression in human tissues, using methods described by, among others, Sambrook and colleagues (*supra*). A cDNA probe containing the entire nucleotide sequence of the IL-20 protein (SEQ ID NO:1) is labeled with ^{32}P using the *redi*prime™ DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using a CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for IL-20 mRNA.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) are obtained from Clontech and are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol

number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and films developed according to standard procedures.

Results of Northern blot experiments performed essentially as described above indicate that a major mRNA of approximately 5 kb is detected predominantly in the thymus, and, to a lesser degree, in the adrenal cortex, spleen, pancreas, and only at very low levels in the lymph node, peripheral blood lymphocytes, fetal liver, adrenal medulla, thyroid, small intestine, stomach and heart. A major mRNA of approximately 1 kb, with a minor mRNA at approximately 5 kb was detected in additional experiments in the testis and spinal cord and, to a lesser extent, in bone marrow and small intestine.

Example 5: Effect of IL-20 on the proliferation of HT-29 cells

Human tumor cell lines including breast carcinoma MDA-MB-231, colon cancer HT-29, prostate cancer PC-3 and osteogenic sarcoma MNNG / HOS were obtained from ATCC and cultured in the medium recommended for each cell line by ATCC.

Tumor cells were harvested by trypsinization and seeded in wells of a 96-well plate at 5,000 cells/well in the appropriate growth medium. IL-20 protein (or supernatant) was then added at concentrations from 0 to 10000 ng/ml in basal medium. Taxol at a concentration of 50 ng / ml is used as a positive control. The appropriate buffer (without protein) is utilized as a negative control. The cells were incubated in a final volume of 200 ul for 4-5 days. AlamarBlue was added to each well to a final concentration of 10%. The cells are incubated for 4 hr. Cell viability was measured by reading in a CytoFluor fluorescence reader with excitation at 530nm and emission at 590nm.

Results from initial experiments performed essentially as described above indicate that IL-20 supernatants have stimulatory effects on the growth of HT-29 cells *in vitro*.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of all publications (including patents, patent applications, journal articles, laboratory manuals, books, or other documents) cited herein are hereby incorporated by reference. The hard copy of the sequence listing and the computer readable form of same are both incorporated herein by reference.